

## **PATENT ABSTRACTS OF JAPAN**

(11)Publication number : **10-094395**

(43)Date of publication of application : **14.04.1998**

---

(51)Int.Cl.

**C12N 15/09**

**A23K 1/16**

**C07D311/92**

**C07D493/04**

**C07K 14/315**

**C12N 1/21**

**C12P 17/18**

**// (C12N 1/21**

**C12R 1:465 )**

**(C12P 17/18**

**C12R 1:465 )**

---

(21)Application number : **09-116652**

(71)Applicant : **F HOFFMANN LA ROCHE  
AG**

(22)Date of filing : **07.05.1997**

(72)Inventor : **REEVES CHRISTOPHER  
DESMOND  
SOLIDAY CHARLES  
LESLIE**

---

(30)Priority

Priority number : **96 16753**  
**97 42935**

Priority date : **07.05.1996**  
**04.04.1997**

Priority country : **US**  
**US**

---

(54) **FRENOLICIN GENE CLUSTER**

TELETYPE CENTER	TELETYPE CENTER	TELETYPE CENTER	TELETYPE CENTER	TELETYPE CENTER
TELETYPE CENTER	TELETYPE CENTER	TELETYPE CENTER	TELETYPE CENTER	TELETYPE CENTER
TELETYPE CENTER	TELETYPE CENTER	TELETYPE CENTER	TELETYPE CENTER	TELETYPE CENTER
TELETYPE CENTER	TELETYPE CENTER	TELETYPE CENTER	TELETYPE CENTER	TELETYPE CENTER

**SOLUTION:** This is a new DNA sequence containing a DNA sequence coding for biochemically equivalent variant to a protein coded by a DNA sequence specified by bases at 636th to 2948th position, 2945th to 3916th position, 4020th to 4844th position, 4841th to 6415th position, 6533th to 7183th position, 7344th to 8897th position, 9164th to 10012th position, etc., in the DNA sequence expressed by the formula, and the cluster is useful for improving the yield of the biosynthesis of an antibiotic, frenolicin B, by a microorganism, especially *Streptomyces roseofulvus*. The cluster

is a discharging pump useful for the production of frenolicin and can be obtained by combining genes of butyrate starter synthase, polyketide synthase, hemiketalase, keto-reductase, cyclase/ dehydrase, keto/enoyl reductase, etc.

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-94395

(43) 公開日 平成10年(1998) 4月14日

(51) Int.Cl. <sup>6</sup>	識別記号	F I	
C 1 2 N 15/09	Z N A	C 1 2 N 15/00	Z N A A
A 2 3 K 1/16	3 0 2	A 2 3 K 1/16	3 0 2 B
C 0 7 D 311/92	1 0 1	C 0 7 D 311/92	1 0 1
493/04	1 0 6	493/04	1 0 6 A
C 0 7 K 14/315		C 0 7 K 14/315	
審査請求 未請求 請求項の数12 O L (全 46 頁) 最終頁に続く			

(21) 出願番号	特願平9-116652	(71) 出願人	591003013 エフ・ホフマン-ラ ロシュ アーゲー F. HOFFMANN-LA ROCH E AKTIENGESELLSCHAFT スイス・シーエイチ-4070バーゼル・グレ ンツアーヘルストラツセ124
(22) 出願日	平成9年(1997) 5月7日	(72) 発明者	クリストファー・デスモンド・リーブズ アメリカ合衆国、ワシントン 98072、ウ ッディンビル、エヌイー・23アールディ・ ブレイス 19403
(31) 優先権主張番号	60/016753	(74) 代理人	弁理士 津国 肇 (外1名)
(32) 優先日	1996年5月7日		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/042935		
(32) 優先日	1997年4月4日		
(33) 優先権主張国	米国 (US)		
		最終頁に続く	

(54) 【発明の名称】 フレノリシン遺伝子クラスター

(57) 【要約】

【課題】 微生物、特にストレプトミセス・ロゼオフルバス(*Streptomyces roseofulvus*)により、抗生物質フレノリシンBの生産に関与する遺伝子を同定、単離、および配列決定すること。

【解決手段】 原核生物細胞内でフレノリシンの生合成を触媒するタンパク質をコードする、配列番号: 22を含む単離された遺伝子クラスター、およびフレノリシンの生合成を触媒するタンパク質をコードする、配列番号: 22と十分に機能的に同等な改変された遺伝子クラスターを提供する。

## 【特許請求の範囲】

【請求項1】 少なくとも1つの以下のDNA配列：

- a) 配列番号：22の636位～2948位の塩基
- b) 配列番号：22の2945位～3916位の塩基
- c) 配列番号：22の4020位～4844位の塩基
- d) 配列番号：22の4841位～6415位の塩基
- e) 配列番号：22の6533位～7183位の塩基
- f) 配列番号：22の7344位～8897位の塩基
- g) 配列番号：22の9164位～10012位の塩基
- h) 配列番号：22の10621位～10105位の塩基
- i) 配列番号：22の11628位～10618位の塩基
- k) 配列番号：22の11809位～12066位の塩基
- l) 配列番号：22の13209位～12154位の塩基
- m) 配列番号：22の13409位～14686位の塩基
- n) 配列番号：22の14767位～16047位の塩基
- o) 配列番号：22の16120位～16371位の塩基
- p) 配列番号：22の16935位～16453位の塩基
- q) 配列番号：22の17088位～17903位の塩基
- r) 配列番号：22の17903位～18898位の塩基
- s) 配列番号：22の18895位～19839位の塩基
- t) 配列番号：22の20907位～19990位の塩基
- w) 配列番号：22の22094位～20904位の塩基

もしくはa)～w)に特定されるDNA配列によってコードされるタンパク質の生化学的に同等な変種をコードするDNA配列、を含むDNA配列。

【請求項2】 a)～w)に特定される全てのDNA配列を含み、好ましくは配列番号：22と同様に並んでいるか、あるいはa)～w)に特定されるDNA配列の少なくとも1つが、生化学的に同等な変種のタンパク質をコードするDNA配列によって置換されている、請求項1に記載のDNA配列。

【請求項3】 前記配列番号：22のDNA配列もしくは生化学的に同等な変種であるタンパク質をコードするDNA配列を含む、請求項2に記載のDNA配列。

【請求項4】 作動可能に発現制御配列に連結した請求項1から3のいずれか1項に記載のDNA配列を含むベクター。

【請求項5】 請求項4に記載のベクターによって形質転換された宿主細胞。

【請求項6】 前記宿主細胞がストレプトミセス属に属する、請求項5に記載の形質転換細胞。

【請求項7】 前記宿主細胞がストレプトミセス・ロゼオフルバス(*Streptomyces roseofulvus*)である、請求項6に記載の形質転換細胞。

【請求項8】 請求項1のa)～w)に規定されたDNA配列によってコードされるか、あるいは配列番号：1～21のアミノ酸配列を有するタンパク質およびそれらの生化学的に同等の変種。

【請求項9】 フレノリシンもしくはそのようなフレノリシンの生合成中間体を調製する方法であって、請求項5～7のいずれか1項に記載の細胞を、適切な培養条件下で培養し、そして該培養物もしくは該細胞から該フレノリシンもしくはその生合成中間体を単離することを特徴とする、調製方法。

【請求項10】 請求項9に記載の方法によって取得される生合成中間体を、当該分野で公知の化学的もしくは他の方法により、そのようなフレノリシンに変換することを特徴とする、フレノリシンの調製方法。

【請求項11】 請求項9または10の方法にしたがって得た前記フレノリシンもしくはフレノリシンの混合物をフレノリシンBに酸化するフレノリシンBの調製方法。

【請求項12】 請求項9～11のいずれか1項に記載の方法を行って、得られた前記フレノリシンを他の飼料組成物成分と混合することを特徴とする、飼料組成物の調製方法。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は、微生物、特にストレプトミセス・ロゼオフルバス(*Streptomyces roseofulvus*)による抗生物質フレノリシンBの生産に関与する遺伝子の同定、単離、および配列決定に関する。

【0002】ポリケチドは、大きな、構造的に多様なファミリーを持つ天然の産物であり、抗生物質特性および薬理学的特性を含む広範囲の生物学的活性を示す。例えば、ポリケチドは、抗生物質であるテトラサイクリンおよびエリスロマイシン、抗ガン剤ダウノマイシン、免疫抑制物質FK506およびラパマイシン、およびフレノリシン、モネンシン、およびアベルメクチンのような家畜用製品によって代表される。

【0003】ポリケチドシンターゼ(PKS)は、脂肪酸シンターゼ(FAS)に関連した多機能酵素である。PKSは、アシルチオエステル、通常、アセチル、プロピオニル、マロニルもしくはメチルマロニル間のクライゼン縮合の反復によりポリケチドの生合成を触媒する。各縮合に続いて、それらは、成長するポリケチド鎖のβ-ケト基におけるケト還元、脱水、およびエノイル還元

を含む還元的サイクルの全てもしくは一部を触媒することにより、または何も触媒しないことにより、生成物に構造的多様性を誘導する。炭素鎖がそれぞれの特定の生成物に特徴的な長さにまで成長した後、炭素鎖はチオール開裂もしくはアシル転移によりシンターゼから放出される。したがって、PKSは一緒に作用して所定のポリケチドを生成する酵素ファミリーからなる。一般に、各PKSにプログラムされている鎖長、鎖構築単位を選択、および還元サイクルの制御された変化は、天然に生じるポリケチド間に見られる変化の一因となっている。

【0004】ストレプトミセスは芳香族ポリケチドの放線菌生産体である：ストレプトミセスにおいては、PKS仲介性合成は、少なくとも3つのPKSオープンリーディングフレーム(ORF)の産物に依存していることが知られている。ORF1は、ケトシンターゼ(KS)およびアシルトランスフェラーゼ(AT)の活性部位をコードし；ORF2は、ORF1の産物に類似しているがKSおよびATモチーフを欠いているタンパク質をコードし；そしてORF3は、別のアシルキャリアタンパク質(ACP)をコードしている。

【0005】

【従来の技術】フレノリシンBは、国内の動物および鳥類に抗コクシジウム物質(anticoccidial)として広く用いられているナフトキノンポリケチドである。Iwaiらは、J. Antibiotics, 31:10, 959-965(1978)において、フレノリシン群の2つの抗生物質、AM-3867IおよびIIを、土壌から単離したストレプトミセス・ロゼオフルバス(*Streptomyces roseofulvus*) AM-3867株の発酵ブロスから単離したことを報告した。AM-3867Iは、その後フレノリシンBと名付けられた新規な抗生物質であることが見いだされ、一方後者はデオキシフレノリシンと同定された。

【0006】BibbらはGene, 142, 31-39(1994)において、10.2kbのDNAフラグメントのストレプトミセス・ロゼオフルバスからのクローニングを報告した。このフラグメントは、フレノリシンおよびナナオマイシンの生成を決定することが予想されるポリケチドシンターゼ(PKS)をコードする遺伝子(fren)を含んでいる。このDNAの連続した5530bpのセグメントを配列決定した。配列の分析により、一つの方向に転写される5つの完全なオープンリーディングフレーム(ORF)(ORF1、2、3、5、4)およびORF3とORF5との間に位置して逆方向に転写されるオープンリーディングフレーム(ORFX)が明らかにされた。ORF1、2、3、4、および5の推定アミノ酸配列は、他のストレプトミセス種のII型PKSの公知の成分の配列に非常に類似している：それぞれ、推定されるヘテロ二量体(ORF1+2)ケトシンターゼ、アシルキャリアタンパク質、シクラーゼおよびケトレダクターゼである。ORF4産物のN-末端およびC-末端半分の類似性も

また、他のイソクロマネキノン抗生物質生産体由来の対応する遺伝子において発見され、シクラーゼをコードする遺伝子の可能な起源を二重に示唆する。ORFXは、frenクラスターだけでなくストレプトミセス種の芳香族抗生物質合成遺伝子の他のクラスターにも存在する未知の機能の新規なクラスの遺伝子を示すようである。fren遺伝子が実際にフレノリシンPKSをコードしていることを証明するために、遺伝子を破壊することによって、これらの遺伝子をAK24158株に導入しようとする試みは、失敗したことが報告されている。

【0007】

【発明が解決しようとする課題】本発明の目的は、フレノリシンをコードする遺伝子クラスターの全長配列を提供し、そのクラスターのオープンリーディングフレームを単離および同定することである。

【0008】また、本発明の目的は、上記単離された任意のもしくは全てのフレノリシンPKS遺伝子で形質転換した操作された宿主細胞を提供することである。

【0009】さらに本発明の目的は、フレノリシンの合成生産の収率を改善することである。

【0010】

【課題を解決するための手段】本発明は、原核生物細胞内でフレノリシンの生合成を触媒するタンパク質をコードする、配列番号：22を含む単離された遺伝子クラスター、およびフレノリシンの生合成を触媒するタンパク質を同様にコードし、配列番号：22と十分に機能的に生化学的に同等な配列を有する改変された遺伝子クラスターを提供する。また、発現制御配列に作動可能に連結した上記遺伝子クラスターを含むベクター、およびそのようなベクターで形質転換された宿主細胞も含まれる。

【0011】また、フレノリシン合成の特定の局面を担う5つの単離された遺伝子サブクラスターも提供される：

配列番号：1、2、3、4および6に示すタンパク質、ならびにそのタンパク質の生化学的に同等な突然変異タンパク質をコードする遺伝子を含む、フレノリシン生産に有用な流出ポンプをコードする単離された第1の遺伝子サブクラスター；

配列番号：8、9、10および11に示すタンパク質、ならびにそのタンパク質の生化学的に同等な突然変異タンパク質をコードする遺伝子を含む、フレノリシン生産に有用なブチレートスターターシンターゼ(butyrate starter synthase)をコードする単離された第2の遺伝子サブクラスター；

配列番号：12、13および14に示すタンパク質、ならびにそのタンパク質の生化学的に同等な突然変異タンパク質をコードする遺伝子を含む、フレノリシン生産に有用なポリケチドシンターゼをコードする単離された第3の遺伝子サブクラスター；

配列番号：15、16、17および18に示すタンパク

質、ならびにそのタンパク質の生化学的に同等な突然変異タンパク質をコードする遺伝子を含む、フレノリシン生産に有用なヘミケタラーゼ、ケトレダクターゼ、およびシクラーゼ／デヒドラーゼをコードする単離された第4の遺伝子サブクラスター；および

配列番号：19および20に示すタンパク質、ならびにそのタンパク質の生化学的に同等な突然変異タンパク質をコードする遺伝子を含む、フレノリシン生産に有用なケト／エノイルレダクターゼおよびヒドロキシラーゼをコードする単離された第5の遺伝子サブクラスター。

【0012】別の実施態様においては、各遺伝子は発現制御配列に作動可能に連結したベクターに独立して挿入されてもよい。

【0013】別の実施態様においては、a) 宿主細胞を配列番号：22の遺伝子クラスターをコードする組換えベクターによって形質転換し、b) そのベクターをその宿主細胞内で発現させ、そしてc) それから産生されたフレノリシンを単離することによるフレノリシンの生合成方法が提供される。

【0014】また、a) 宿主細胞を配列番号：22の遺伝子クラスターをコードする組換えベクターによって形質転換し、b) そのベクターをその宿主細胞内で発現させ、c) それから産生されたフレノリシンを単離し；そしてd) フレノリシンの混合物を酸化して実質的に全ての単離生成物をフレノリシンBに変換することによる抗生物質フレノリシンBの組換え生合成方法も提供される。

【0015】また、配列番号：1～20の単離されたタンパク質およびそれらの生化学的に同等な変種ならびにそのようなタンパク質をコードする遺伝子も提供される。

【0016】本発明のさらなる目的は、少なくとも1つの以下のDNA配列を含むDNA配列を提供することである：

- a) 配列番号：22の636位～2948位の塩基
- b) 配列番号：22の2945位～3916位の塩基
- c) 配列番号：22の4020位～4844位の塩基
- d) 配列番号：22の4841位～6415位の塩基
- e) 配列番号：22の6533位～7183位の塩基
- f) 配列番号：22の7344位～8897位の塩基
- g) 配列番号：22の9164位～10012位の塩基
- h) 配列番号：22の10621位～10105位の塩基
- i) 配列番号：22の11628位～10618位の塩基
- k) 配列番号：22の11809位～12066位の塩基
- l) 配列番号：22の13209位～12154位の塩基
- m) 配列番号：22の13409位～14686位の塩基

基

n) 配列番号：22の14767位～16047位の塩基

o) 配列番号：22の16120位～16371位の塩基

p) 配列番号：22の16935位～16453位の塩基

q) 配列番号：22の17088位～17903位の塩基

r) 配列番号：22の17903位～18898位の塩基

s) 配列番号：22の18895位～19839位の塩基

t) 配列番号：22の20907位～19990位の塩基

w) 配列番号：22の22094位～20904位の塩基

もしくはa)～w)に特定されるDNA配列によってコードされるタンパク質の生化学的に同等な変種をコードするDNA配列、さらにより特定するとa)～w)に特定される全てのDNA配列を含むDNA配列であって、好ましくは配列番号：22と同様に並んでいるDNA配列、あるいはa)～w)に特定されるDNA配列の少なくとも1つが、生化学的に同等な変種のタンパク質をコードするDNA配列によって置換されているDNA配列、さらにより特定すると、配列番号：22のDNA配列を含むDNA配列もしくは生化学的に同等な変種のタンパク質をコードするDNA配列。

【0017】本発明のさらなる目的は、発現制御配列に作動可能に連結したこのようなDNA配列を含むベクター、そのようなベクターで形質転換した宿主細胞を提供することである。特に、この場合、この宿主細胞はストレプトミセス属の一員、例えばストレプトミセス・ロゼオフルバス (*Streptomyces roseofulvus*) である。

【0018】本発明の目的はまた、上記a)～w)に規定されるDNA配列によってコードされるタンパク質もしくは配列番号：1～21に与えられるアミノ酸配列を有するタンパク質、またはそれらの生化学的に同等な変種を提供すること、ならびにフレノリシンもしくはそのようなフレノリシンの生合成中間体の調製方法を提供することである。この調製方法は、上記に特定した細胞を適切な培養条件下で培養し、その培養物もしくはその細胞からフレノリシンを単離することを特徴とし、そしてこのフレノリシン調製方法は、上記に特定した方法によって得た生合成中間体を当該分野で公知の化学的もしくは他の方法によってそのようなフレノリシンに変換することを特徴とし、そしてフレノリシンBの調製方法は、上記の方法にしたがって得たフレノリシンもしくはフレノリシンの混合物をフレノリシンBに酸化するものである。

【0019】さらに本発明の目的は、上記に特定された方法を実施して得られたフレノリシンを別の飼料組成物成分と混合することを特徴とする飼料組成物の調製方法を提供することである。

#### 【0020】

【発明の実施の形態】以下の規定は、本明細書中の発明を記載するために用いられる種々の用語の意味および範囲を説明および定義するために示される。

【0021】「宿主細胞」は、DNA配列、特に本発明のフレノリシン遺伝子クラスターを含む組換えベクターのレシピエントとして用いられ得る原核細胞微生物に由来する細胞である。この用語は、形質転換されたものとの細胞の子孫を含む。単一の親細胞の子孫は、偶発的もしくは計画的な変異により、形態学的にあるいはゲノムもしくは全DNA相補物においてもとの親と必ずしも完全に一致する必要はない。関連する性質、例えば、特定のDNA配列の存在によって特徴づけられた親と十分に類似している親細胞の子孫は、この規定に含まれる。

【0022】用語「異種の」が、コード配列および制御配列のような核酸配列に関連する場合には、この用語は、通常は組換え構築物の領域に付随しない配列、および／または通常は特定の細胞に付随しない配列を意味する。したがって、核酸構築物の「異種の」領域とは、天然では別の分子と一緒に見いだされることはない、もう一つの核酸分子内にあるかそれに付着している同定し得る核酸セグメントである。例えば、構築物の異種領域は、天然ではそのコード領域と一緒に見いだされない配列が近接したコード配列を含み得る。異種コード配列の別の例は、コード領域自身が天然には見いだされない構築物である（例えば、天然の遺伝子とは異なるコドンを含む合成配列）。同様に、通常その宿主細胞中に存在しない構築物で形質転換された宿主細胞は、本発明の目的のために異種とみなされる。アレル変異もしくは天然に生じる変異事象は、本明細書中で用いられるような異種DNAのもととはならない。

【0023】「コード配列」は、適切な調節配列の制御下におかれるとin vitroもしくはinvivoで転写されるか（DNAの場合）翻訳される（mRNAの場合）核酸配列である。コード配列の境界は5'（アミノ）末端の翻訳開始コドンおよび3'（カルボキシ）末端の翻訳停止コドンによって決定される。翻訳終結配列は通常コード配列の3'に位置する。

【0024】「核酸配列」は、天然のヌクレオシド塩基であるアデニン、グアニン、シトシン、チミンおよびウラシルを含む、原核生物および真核生物のmRNA、原核生物および真核生物のmRNA由来のcDNA、ゲノムDNA、および合成DNAおよびRNA配列を含むがこれに限定されない。この用語はまた、4-アセチルシトシン、8-ヒドロキシー-N6-メチルアデニン、アジリジニルシトシン、シュードイソシトシン、5-（カル

ボキシヒトロキシルメチル）ウラシル、5-フルオロウラシル、5-ブロモウラシル、5-カルボキシメチルアミノメチル-2-チオウラシル、5-カルボキシメチルアミノメチルウラシル、ジヒドロウラシル、イノシン、N6-イソペンテニル-アデニン、1-メチルアデニン、1-メチルシュードウラシル、1-メチルグアニン、1-メチル-イノシン、2, 2-ジメチルグアニン、2-メチルアデニン、2-メチルグアニン、3-メチル-シトシン、5-メチルシトシン、N6-メチルアデニン、7-メチルグアニン、5-メチルアミノメチルウラシル、5-メトキシアミノメチル-2-チオウラシル、 $\beta$ -D-マンノシルクエオシン、5'-メトキシカルボニルメチルウラシル、5-メトキシウラシル、2-メチルチオ-N6-イソペンテニルアデニン、ウラシル-5-オキシ酢酸メチルエステル、ウラシル-5-オキシ酢酸、オキシブトキソシン、シュードウラシル、クエオシン、2-チオシトシン、5-メチル-2-チオウラシル、2-チオウラシル、4-チオウラシル、5-メチルウラシル、および2, 6-ジアミノプリンを含むがこれらの限定されない、1以上の他の塩基を有する配列も包含する。

【0025】核酸「制御配列」とは、決められた宿主細胞内で所定のコード配列の転写および翻訳に必要なかつ十分な、翻訳開始コドンおよび翻訳停止コドン、プロモーター配列、リボソーム結合部位、ポリアデニル化シグナル、転写終結配列、上流調節ドメイン、エンハンサーなどをいう。所望の遺伝子の転写および翻訳に必要なかつ十分な制御配列が存在する限り、これらの制御配列の全てが組換えベクター内に存在する必要はない。

【0026】「作動可能に(operably)もしくは機能的に(operatively)連結した」とは、コード配列と制御配列が所望の機能を果たすような配置をいう。したがって、コード配列に作動可能に連結した制御配列は、そのコード配列の発現を行うことができる。コード配列は、DNAポリメラーゼがプロモーター配列に結合して、コードされたタンパク質に翻訳され得るmRNAにコード配列を転写する場合、細胞内で転写調節領域に作動可能に連結するかもしれないその制御下にある。この制御配列は、それらがコード領域の発現を行う限り、コード領域に連続している必要はない。したがって、例えば、翻訳されていないが転写されている配列が、プロモーター配列とコード配列との間に介在してもよく、このプロモーター配列はそれでもコード配列に「作動可能に連結している」と考えてよい。

【0027】「選択マーカー」によって任意の遺伝的マーカーが意味され、このマーカーは、このマーカーをそのゲノムに有している細胞の集団を選択するために用いることができる。選択マーカーの例には以下の物が含まれる：栄養要求性マーカー。これにより、栄養素もしくは補充物（例えば、チミジン、ジアミノピメリン酸もし

くはビオチン)を含むか含まない最少培地上で細胞が生育し得る能力により細胞を選択する;代謝マーカー。これにより、唯一の炭素源として適切な糖を含む最少培地上で細胞が生育し得る能力、もしくは細胞が適切な染料もしくは色素基質を有する色の付いたコロニーを形成する能力により細胞を選択する、および;薬剤耐性マーカー。これにより、1以上の適切な薬剤(例えばテトラサイクリン、アンピシリン、カナマイシン、ストレプトマイシンもしくはナリジキシン酸)を含む培地上で細胞が生育する能力により細胞を選択する。

【0028】核酸構築物は、一般に、転写カセットとして提供される。場合によって1つのイントロンがその構築物に含まれてよく、好ましくは3100bpでそのコード配列の5'側に位置してもよい。一般に、この構築物は宿主細胞のゲノムに組み込まれるようになるのではなく、この構築物は組み込まれない発現カセットの一部として宿主に導入されることが好ましい。

【0029】本発明において用いるための構築物は、ベクター、転写カセット、翻訳カセットおよびプラスミドを含む。転写および翻訳開始配列は、好ましくは、転写開始調節領域(時折「プロモーター」ともいわれる)およびリボソームがmRNAに結合して開始コドンで翻訳を開始することを可能にする「リボソーム結合部位」を含む5'-非翻訳配列の翻訳開始調節領域を含む。開始制御領域の全ての転写および翻訳機能性要素は、同じ遺伝子から誘導されるか入手し得ることが好ましい。いくつかの実施態様においては、プロモーターは配列の付加または欠失によって、あるいは別の配列による置換によって改変し得る。「入手し得る」によって、目的DNA配列の転写および翻訳の所望の特異性を与える天然の遺伝子の配列に十分類似した調節配列が意図される。それは、天然および合成の配列、ならびに合成と天然の配列の組合せであってもよい配列を含む。転写開始領域、もしくはプロモーター要素には、それが目的のDNA配列の所望の転写レベルを提供する限り、任意の領域を用いることができる。この転写開始領域は宿主細胞に対して、および/または転写されるべきDNA配列に対して固有のものであるかまたは同種であってもよく、あるいは宿主細胞に対して、および/または転写されるべきDNA配列に対して外来であるかまたは異種であってもよい。宿主細胞に対して外来であるとは、転写開始領域を含む構築物が挿入される宿主には、その転写開始領域は見いだされないということを意図する。DNA配列に対して外来であるとは、通常は目的のDNA配列には伴われていない転写開始領域を意図する。

【0030】転写開始調節領域の下流もしくはその制御下には、宿主の遺伝子型に1以上の変更を与えるかもしくは宿主の表現形を調整する目的の核酸配列を挿入するためのマルチプルクローニング部位がある。都合のよいことに、このマルチプルクローニング部位は効果的な方

法で種々の核酸配列のために用いることができる。クローニング部位に挿入される核酸配列は、コード配列が目的のポリペプチドをコードする場合に、それが、適切なmRNA分子の生産を阻止するかおよび/または変形のスプライシングをされたか、または、通常ではないmRNA分子の生産をすることができる潜在のスプライス部位を欠除している限り、目的のポリペプチドをコードする任意のオープンリーディングフレームを有していてもよい。この核酸配列はcDNAであってもよい;それはまたゲノム配列に相補的な配列であってもよく、ここでこのゲノム配列は1以上のオープンリーディングフレーム、イントロン、非コード化リーダー配列であり得、あるいは上記相補的な配列が、転写、メッセンジャーRNAのプロセッシング(例えばスプライシング)もしくは翻訳を阻害する任意の配列であってもよい。

【0031】主に用いられる終結領域は、好都合なものであってよい。なぜなら、終結領域は比較的交流可能なようだからである。この終結領域は意図した目的の核酸配列に対して固有のものであってよいし、あるいは別の起源に由来してもよい。

【0032】転写ベクターは、特定のコード配列が適切な調節配列と共にベクター内に位置するように構築され、このコード配列の制御配列に対する位置および方向は、このコード配列がこの制御配列の「制御」下で転写されるようになっている。目的の特定のタンパク質をコードする配列の改変は、この目的を達成するのに好ましい。例えば、いくつかの場合には、上記の配列が適切な方向で上記の制御配列に付着できるようにその配列を改変すること;あるいはリーディングフレームを維持することも必要となり得る。この制御配列および他の調節配列は、ベクターに挿入する前にコード配列と連結させてもよい。あるいは、このコード配列を、既に制御配列、およびリーディングフレーム内にその制御配列を伴ってかつその調節制御下にある適切な制限部位を有している発現ベクターに直接クローニングすることもできる。

【0033】用語「生化学的に同等な変種」は、いくつかの点で本明細書中に開示されている特定の配列とは異なるにもかかわらず、同じもしくは実質的に同じ機能性を示すタンパク質もしくは核酸配列を意味する。例えば、cDNAの場合、これは、具体的に開示された核酸配列ではない核酸配列を含む改変された配列は、cDNAが、次に本発明のタンパク質をコードするmRNAをコードするかぎり、この代わりの配列が含まれることを意味する。このような改変はわずか2、3個の、もしくは多くの核酸の置換を含み得る。この改変は縮重したコード配列の置換、もしくはあるコード配列の別の配列による置換を含み得る;非天然の核酸の導入も含まれる。機能的な基準を満たしている限り、変更したDNAが本明細書中に開示されているDNAと必ずしもハイブリダイズする必要はない。同様に、本発明のタンパク質の場



合、機能的に影響を及ぼさないアミノ酸配列の変更を行ってもよい。このような「生化学的に同等な突然変異タンパク質」は、1つのアミノ酸のもう一つのアミノ酸での置換、側鎖を改変したもしくは非天然のアミノ酸の使用、および短縮化を含み得る。当業者は、どの部位が基本的な機能に影響を与えない変更にも最も適しているかを認識し得るであろう。そのような変更した配列のほとんどは、本明細書中に開示した配列と、少なくとも50%、好ましくは少なくとも75%、そして最も好ましくは少なくとも90%の相同性を示すであろうことが期待される。

【0034】24 kbのフレノリシンクラスターを配列決定するための手順は、大規模なゲノム配列決定に用いられる方法から適合させたものである。クラスターの部分をランダムに400~1800 bpのフラグメントに消化し、そこから自動配列決定用にライブラリーを構築した。配列データを収集し、Macintosh Quandra 上でDNASTARソフトウェアを用いて操作した。

【0035】制限地図(図1)に見られるように、この24 kbフレノリシンクラスターは、好都合にもNotIにより3つの大きなフラグメント(7.5、9.0および9.5 kb)および1つの小さなフラグメント(1.5 kb)に断片化された。小さい方のNotIフラグメントはフレノリシン生成遺伝子クラスターのPKS領域の中央に存在し、その両側に大きなNotIフラグメントが近接している。図1で左から右に並んだ各々のNotIフラグメントを、pBluescript にサブクローニングし、それぞれNot6、Not7、Not2、およびNot3と命名したクローンを得た。ショットガン配列決定法のために、3つの大きなNotIフラグメントを用いてMspIフラグメントの別々のライブラリーを生成した。NotIフラグメントの3' および5' 末端の配列決定により、「anchor」配列が提供され、配列データとマップ化されたクラスターとの整列化が容易になった。MspIによる部分消化の後、各NotIサブクローンから0.4~1.8 kbのフラグメントをゲル単離することによりランダムライブラリーを構築した。これらのフラグメントをpBluescript のAccI部位に連結し、適切なE.coli宿主(DH5αもしくはXL1-Blue)内に形質転換して各ライブラリーを作成した。これらのライブラリーを播種し十分な数のサブクローンを選択して二本鎖配列決定のための鋳型を提供した。

【0036】この鋳型の配列決定は、ポリメラーゼ連鎖反応および蛍光的に標識したヌクレオチドを用いるApplied Biosystems(Foster City, CA)のTaq dye ーデオキシターミネーターサイクル配列決定反応により行った。配列決定ゲルを読み取り、データを収集し、そして配列をABI 373A自動シーケンサーによって分析した。配列データを得るにしたがって、それを以前に決定された配列と並べて、オーバーラップし、連続した配列の領域(整列群: contigs)を形成した。配列データが蓄積するにつ

れ、この整列群はより大きな整列群に合流した。ほとんどの領域では、両方の鎖が3、4回配列決定され、最終的に全クラスターの信頼し得る配列を得た。

【0037】遺伝子はこの配列内に、そのコドンの第3位に高頻度でGまたはCのいずれかを有するオープンリーディングフレーム(ORF)として存在したが、このことはストレプトミセス遺伝子の一般的な特性である(Bibb MJ, et al., Gene 30:157(1984))。このように存在する遺伝子のアミノ酸配列をSWISS-PROTデータベース内の配列と比較した。SWISS-PROTデータベースが特定の配列を欠いているような場合には、予想される遺伝子のヌクレオチド配列もGenBank ストレプトミセスDNA配列と比較した。このように、フレノリシンクラスター内の遺伝子に類似した配列を有する遺伝子を見だし、推定される機能を当てはめた。

【0038】フレノリシンクラスターの完全な配列を配列番号: 22に供した。クラスターの左端から、各遺伝子について以下に議論を始める。フレノリシン生成の仮定的な経路を、その経路中で各遺伝子が果たす仮定的役割とともに図2に示した。示した遺伝子の機能を表1にまとめた。クラスター内の遺伝子の順番は、生成工程の順番と同じようである。この特徴はマクロライド経路について以前に報告されているが(Donadio S, et al., 1991, Science 252:675)、芳香族ポリケチド経路については報告されていない。

【0039】遺伝子A(配列番号: 1のタンパク質をコードする配列番号: 22の636位~2948位の塩基)は、クラスター内の最初の遺伝子と考えられている。なぜなら、その上流に1 kbの非コード領域があり、その領域の上流の部分遺伝子の5' 端は配列データベースのどの配列とも相同ではないからである。遺伝子A~DおよびFはサブクラスターを含み、これはフレノリシンの放出に関与しているようである。遺伝子A~Dの産物は、ABCトランスポーターシステムの構成成分を示し得る。ABCトランスポーターの特徴的な性質は2つのATP結合ドメインを含むことであり、その各々が2つの保存された配列モチーフ、および2つの膜ドメインを有し、これら膜ドメインの各々が6つの膜貫通配列を有する(Higgins CF, Ann Rev Cell Biol 8:67(1992))。この2つのATP結合ドメインおよび2つの膜ドメインは、真核生物多剤耐性タンパク質の場合(この場合、ポリペプチドは2つの半分からなり、それらは内部の重複を示す(Chen C et al., 1986, Cell 47:381))のように、全て同じポリペプチド上にあり得る。細菌のシステムはしばしば別々のポリペプチド上に異なるドメインを有するが、どの場合も膜成分およびATP結合成分は多量体複合体に組み立てられ、その結果ドメインの三次元的位置取りは全てのトランスポーター内で同じである。細菌の取り込みシステムは、一般に膜の外側にさらなる成分を有しており、これは輸送される基質に結合する; 放出

システムはこの成分を欠いているようである(Reizer J, et al., Prot Sci 1:1326(1992))。

【0040】遺伝子AおよびB(配列番号:2のタンパク質をコードする配列番号:22の2945位~3916位の塩基)、ならびに遺伝子C(配列番号:3のタンパク質をコードする配列番号:22の4020位~4844位の塩基)および遺伝子D(配列番号:4のタンパク質をコードする配列番号:22の4841位~6415位の塩基)は、移行的(translationally)にカップリングし(1つの遺伝子の停止コドンが隣の遺伝子の開始コドンに隣接しているという意味)、正確な化学量論によって同時発現される遺伝子にはよくあることである。遺伝子Dは、各々2つの保存された配列モチーフを有する2つのATP結合ドメインを含む可溶性の524アミノ酸タンパク質をコードする。この遺伝子Dタンパク質は、2つの相同な半分からなり、その各々が*S. peucetius*の推定される放出システムの可溶性成分、DrrB(Guilfoyle PG & Hutchinson CR, Proc Natl Acad Sci 88:8553(1991))の相同性を有する。このDrrBは、アントラサイクリン類、ドキソルビシンおよびダウノルビシンを細胞からくみ出すことが提案されている。この遺伝子D産物はまた、多くの生物の多剤耐性ポンプに相同性がある。遺伝子Cは274アミノ酸のタンパク質をコードし、そのようなタンパク質に典型的な6つの推定膜貫通ドメインを有し、いくつかの細菌ペプチドパーミターゼの2つの膜性分のうち1つに対して良好な類似性を有する。遺伝子Cによってコードされるこのタンパク質はまた、結合タンパク質依存性取り込みシステムの膜成分にしばしば見られるモチーフを含んでいる。遺伝子Bは、遺伝子Aに移行的にカップリングしており、また、6つの推定される膜貫通配列を有するタンパク質をコードし、そして上記の同じ細菌ペプチドパーミターゼの他の膜成分と配列類似性を共有する。遺伝子AおよびBが移行的にカップリングしているので、遺伝子Aが輸送にある役割を果たすことが可能と思われる;しかし、遺伝子Aは、配列データベース中の他の遺伝子と顕著な相同性を示さない。

【0041】遺伝子F(配列番号:6のタンパク質をコードする配列番号:22の7344位~8897位の塩基)は、アクチノホルジンactVA-ORF1遺伝子(Caballero JL, et al., Mol Gen Genet 230:401(1991))と相同であり、actII-ORF2遺伝子(Fernandez-Moreno MA, et al., Cell 66:769(1991))との相同性は幾分低い。これらの遺伝子はどちらもアクチノホルジン放出に関与することが提案されてきた。遺伝子Fはまた*Nocardia lacamdurans*のセファマイシンCクラスターのcmcT遺伝子(Coque JR, et al., EMBO J 12:631(1993))に類似している。したがって、遺伝子Fは、遺伝子A、B、CおよびDと協調して作用し得るか、あるいはATP加水分解によって直接的に行われることのない独立した放出システムを示

し得る。遺伝子Dと遺伝子Fとの間には遺伝子E(配列番号:5のタンパク質をコードする配列番号:22の6533位~7183位の塩基)があり、それはデータベース内の他の遺伝子と顕著な類似性を持たない216アミノ酸の可溶性タンパク質をコードする。

【0042】遺伝子G(配列番号:7のタンパク質をコードする配列番号:22の9164位~10012位の塩基)は、他のポリケチド遺伝子クラスター内の転写アクチベーター〔アクチノホルジンクラスター内のactII-ORF4(Fernandez-Moreno MA, et al., Cell 66:769(1991))およびダウノルビシンクラスター内のdnrI(Stutzman-Engwall KJ, et al., J Bact 174:144(1992))を含む〕に顕著な類似性を有するタンパク質をコードする。それはおそらくクラスター内の多くの他の遺伝子を、以下に検討する推定プロモーターモチーフとの相互作用を介して活性化する。

【0043】遺伝子H、I、JおよびKのセットは全て、おそらくジケチドスターユニット(例えば、ブチリル-ACP、クロトニル-ACP、ヒドロキシブチリル-ACP、もしくはアセトアセチル-ACP)の合成、およびフレノリシン合成を開始するためのそのポリケチドシンターゼへの移行に関与する1つのサブクラスター内に存在する。遺伝子I(配列番号:9)タンパク質は、*E. coli* fabH タンパク質と著しい相同性を有する。これはアセトアセチル-ACPを形成するためのアセチル-CoAとマロニル-CoAの縮合を触媒し、II型システムにおける脂肪酸合成を開始すると考えられている(Tsay J-T, et al., 1992, J Biol Chem 267:6807)。fabHと弱い類似性を持つ遺伝子がダウノマイシンの(Ye J, et al., 1994, J Bacteriol 176:6270) およびドキソルビシンの(Grimm A, et al., 1994, Gene 151:1)遺伝子クラスター内に見られるが、これが、活性部位のシステインの周りに保存された配列モチーフを有する芳香族ポリケチド遺伝子クラスター内に見いだされた最初のfabHホモログである。fabHホモログはアクチノホルジンクラスター内には存在しない。遺伝子Mは、異種システムにおいて他のポリケチドシンターゼ遺伝子と一緒に同時発現されると、自由な鎖長特異性を与えるように思われる(McDaniels R, et al., 1993, J Am Chem Soc 115:11671)が、本明細書で検討される*S. roseofulvus*の株は、18炭素フレノリシンのみを産生し、ナノオマイシンもしくはカラファンジン(kalafungin)のような16炭素構造は産生しない。したがって、遺伝子I(配列番号:9のタンパク質をコードする配列番号:22の11628位~10618位の塩基)は、フレノリシン合成の開始および18炭素長を確保することにおいて鍵となる役割を果たし得る。遺伝子Iに移行的にカップリングしているのが遺伝子H(配列番号:8のタンパク質をコードする配列番号:22の10621位~10105位の塩基)であり、種々のオキシドレダクターゼ遺伝子

と、およびアクチノルホジnkラスタのactI-ORF2の領域と弱い類似性を有する。遺伝子Hの産物はアセトアセチル-ACPのブチリル-ACPへの還元に関与し得る。

【0044】遺伝子J（配列番号：10のタンパク質をコードする配列番号：22の11809位～12066位の塩基）は、遺伝子N（下記に示す）のそれとは異なるアシルキャリアタンパク質をコードし、スターターユニットを形成するために遺伝子Iタンパク質（遺伝子Hタンパク質と一緒に）によって用いられ得る。遺伝子K（配列番号：11のタンパク質をコードする配列番号：22の13209位～12154位の塩基）のタンパク質は、ドキシロビシンの遺伝子クラスター(Grimm A, et al., 1994, Gene 151:1) およびダウノマイシンの遺伝子クラスター(Ye J, et al., 1994, J Bacteriol 176:6270)由来の2つの推定アシルトランスフェラーゼに相同である。それらは全てアシルトランスフェラーゼに典型的な活性部位モチーフ(GlyHisSer)を含む。アクチノルホジnk遺伝子クラスターは、遺伝子IもしくはKのいずれのホモログも含まない。

【0045】遺伝子L、MおよびNは、コアポリケチドシンターゼ複合体をコードし、アクチノルホジnkラスタのactI遺伝子(Fernandez-Moreno MA, et al., 1992, J Biol Chem 267:19278)と相同である。遺伝子L（配列番号：12のタンパク質をコードする配列番号：22の13409位～14686位の塩基）タンパク質は、 $\beta$ -ケトアシル合成およびアシル転移の活性部位モチーフを含む。遺伝子M（配列番号：13のタンパク質をコードする配列番号：22の14767位～16047位の塩基）は、遺伝子Lに相同であるが、活性部位モチーフを欠いている。それは機能的ポリケチドシンターゼ複合体に必要とされる。遺伝子N（配列番号：22の16120位～16371位の塩基）は、配列番号：14のタンパク質をコードし、それはアシルキャリアタンパク質である。これらの遺伝子の近接に基づく、ポリケチドシンターゼは遺伝子Nのアシルキャリアタンパク質を用いるが、fabHホモログは遺伝子Jのアシルキャリアタンパク質を用いるように思われる。

【0046】遺伝子O、P、QおよびRは、第4のサブクラスターを含む。遺伝子O（配列番号：15のタンパク質をコードする配列番号：22の16935位～16453位の塩基）は、アクチノルホジnkラスタのactVI-ORF4遺伝子(Fernandez-Moreno MA, et al., 1994, J Biol Chem 269:24854)に相同であり、ヘミケタール構造の形成および脱水を包含する工程を触媒することが提案されている。遺伝子P（配列番号：22の17088位～17903位の塩基）は、配列番号：16のタンパク質をコードし、それはアクチノルホジnkラスタのactIIIケトレダクターゼ(Hallam SE, et al., 1988, Gene 74:305)に相同である。遺伝子Q（配列番号：22の

17903位～18898位の塩基）は、配列番号：17のタンパク質をコードし、それはアクチノルホジnkラスタのactVIIシクラーゼ/デヒドラーゼ(Fernandez-Moreno MA, et al., 1992, J Biol Chem 267:19278)に相同であり、最初の環状化反応を触媒することが提案されている。遺伝子R（配列番号：22の18895位～19839位の塩基）は、配列番号：18のタンパク質をコードし、それはアクチノルホジnkラスタのactIVタンパク質で、第2の環状化反応に関与するシクラーゼ/デヒドラーゼに非常に類似している(Fernandez-Moreno MA, et al., 1992, J Biol Chem 267:19278)。遺伝子Rは、遺伝子Qに移行的にカップリングしており、そして遺伝子Qは遺伝子Pに移行的にカップリングしているので、これら3つの遺伝子は単一の転写ユニット内にある。

【0047】遺伝子S（配列番号：19のタンパク質をコードする配列番号：22の20907位～19990位の塩基）は、ある領域内のactVA-ORF4、actVI-ORF1およびactVI-ORF2遺伝子に弱い類似性を示す。actVI-ORF1およびactVI-ORF2遺伝子産物は、それぞれ立体特異性ケト還元および立体特異性エノイル還元を触媒することが提案されている(Fernandez-Moreno MA, et al., 1994, J Biol Chem 269:24854)。遺伝子T（配列番号：20のタンパク質をコードする配列番号：22の22094位～20904位の塩基）は、actVA-ORF5(Caballero JL, et al., 1991, Mol Gen Genet 230:401)に非常に類似しており、actVA中間体(Cole SP, et al., 1987, J Antibiot 40:340)のヒドロキシ化をコードしているので、遺伝子Tは、おそらくフレノリシン生合成において類似したヒドロキシラーゼをコードしている。したがって、遺伝子Sは、おそらく上記の立体特異性還元の一方または両方をコードしている。

【0048】推定プロモーターモチーフ（正確に7個の塩基によって分離されている2以上のTGCA配列）は、アクチノルホジnkおよびフレノリシン遺伝子クラスター内のいくつかの推定オペロンの第1のORFの開始コドンの約20～50bp上流にある。イソクロマンキノン生合成遺伝子クラスターの推定プロモーター領域にのみこのモチーフが存在するということは、それがおそらく転写活性化に関与する因子の結合部位を代表するということを示す。このモチーフを含むアクチノルホジnkラスタ内でプロモーターから転写されされる遺伝子は、actII-ORF4遺伝子産物(Gramajo HC, et al., 1993, Mol Microbiol 7:837)によって調節されることが公知である。したがって、フレノリシンラスタ内の遺伝子G（アクチノルホジnkラスタ内のactII-ORF4遺伝子に相同である）は、このモチーフを含むフレノリシンラスタ内のプロモーターの活性化に関与する最もありそうな調節因子である。このモチーフを含むプロモーターの多くは二方向性の転写を与えることができるよう

である。

【0049】クラスターの右端には、種々の生物由来のグリセルアルデヒド-3-リン酸デヒドロゲナーゼ(G3PDH)をコードする遺伝子と非常に類似した遺伝子がある。G3PDH 遺伝子がクラスターの内部に存在することはありそうもないので、これはおそらくクラスターの右側の境界を規定するものである。しかし、*S. arenae*におけるペンタレノラクトン産生の間に、ペンタレノラクトン感受性G3PDH は、その産物に耐性の遺伝的に異なるイソ型のものと置き換えられるということが観察された(Frohlich K-U, et al., 1989, J Bacteriol 171:6696)。したがって、ペンタレノラクトンのように、フレノリシンもまた、GAPDH を阻害し、そして産生の間に耐性のイソ型が産生されることもあり得る。

【0050】フレノリシンクラスターの配列決定された領域の遺伝子で、デオキシフレノリシンエポキシダーゼもしくはフレノリシンBレダクターゼのどちらかをはつきりとコードしているものはない。これらの活性の一方をコードしているかもしれない遺伝子はAおよびEであり、それらはデータベースの何者とも適合しないし、未だに機能を与えられていない。フレノリシンは、ストレプトミセス属および他の細菌に対して毒性であるが、フレノリシンBが最も毒性であり、デオキシフレノリシンはそれより毒性が少なく、そしてフレノリシンエポキシダーゼがさらに毒性が少ない。したがって、フレノリシンBレダクターゼ活性もしくはデオキシフレノリシンエポキシダーゼ活性がフレノリシン遺伝子クラスターから欠けているならば、形質転換体は、得られた産物によって害されるであろう。

【0051】*S. roseofulvus* および他のストレプトミセス属におけるフレノリシンクラスター遺伝子进行操作するために、いくつかの*E. coli*/*Streptomyces* シャトルベクターを構築した。ストレプトミセス内での複製のため、これらのベクターは、プラスミドpSG5由来の温度感受性のレプリコン(Muth G et al., 1988, Mol Gen Genet 211:424; Muth G, et al., 1989, Mol Gen Genet 219:341)を用い、それは*S. ghanaensis*においてゲノム当たり約15のコピー数を有し、そして宿主範囲が広い。このレプリコンの別の利点は、上昇した温度(28℃のかわりに35℃)での形質転換体の生育が、自律的な生育を妨げ、それによりプラスミドの染色体への組み込みについての選択が可能になり、この利点は工業生産用の遺伝的に安定な株を提供するために必要である。*E. coli*におけるこれらのベクターの伸長の基幹は、Litmus28もしくは38(New England Biolabs, Beverly, Mass., USA)のいずれかであり、それらは、pUC ベクターをもとにしている。ストレプトミセスの宿主を形質転換するための一本鎖DNAの使用は、二本鎖型の同じベクターに比べて相同組み込みの頻度が10~100倍高いことを示してきたので(Hillemann D, et al., 1991, Nucleic Acids Res

19:777)、LitmusベクターがM13 複製起点を含むため、Litmusベクターを用いることができ、それにより*E. coli*宿主をM13 ヘルパーファージによって感染させて一本鎖DNAの単離が可能になる。

【0052】本発明のDNA配列、特にフレノリシンクラスター由来の遺伝子(もしくは転写ユニット)を過剰発現させるため、このDNA配列、遺伝子もしくはオペロンを強いストレプトミセスプロモーター、特にtipAおよびermE\* プロモーターにつなぐ(図4)。これを行うため、目的の遺伝子もしくは領域を、所望の制限部位を与えるプライマーを用いてPCRにより増幅し、このPCR産物を適切なベクター、例えば、pErmE もしくはpTipA にクローン化する(図5)。次いで、各プロモーター-遺伝子キメラ構築物を、選択マーカーの2つの半分を再構築する制限部位を用いて(例えば、アンピシリン耐性遺伝子の独特なAseI部位を用いて)、上記のシャトルベクター内に便利なように移す。目的の遺伝子もしくは領域はまた、その天然のプロモーターから、そのプロモーターを改変することによってもしくは上記構築物の多数のコピーを導入することによって達成される高レベルの発現を伴って、発現され得る。任意のこれらの遺伝子の過剰発現が、より高いフレノリシンの力価を導くならば、これらのベクターにより、上昇した温度でプラスミドを維持しつつ一方で選択圧を維持しながら相同的組み込み事象を起こさせることにより安定な生産株を生成することが可能になる。

【0053】ストレプトミセスにおける転写開始の効果的なプロモーター要素は、ermE\* およびtipA、ならびに、高レベルで発現されることが公知の多くの遺伝子に由来するプロモーターを含む。調節配列を、種々の標準的な技術、例えば、所望の部位に独特の制限部位を作成する部位特異的変異(Kunkel TA, 1985, Proc. Natl. Acad. Sci. U.S.A., 82:488-492)、リンカーオリゴヌクレオチドの合成、もしくは所望の制限部位を作成するプライマーを用いるPCRによる増幅、を用いて目的の遺伝子につないでもよい。理想的な実施態様においては、目的のコード配列は、その遺伝子のコドン#1で翻訳開始させるように調節配列につながれている。

【0054】フレノリシンの組換え産生用の宿主細胞は、組換えフレノリシン遺伝子クラスターもしくは本発明の対応するDNA配列を保持することができる任意の生物から誘導することができる。したがって、本発明の宿主細胞は、原核生物もしくは真核生物のいずれに由来することもできる。しかし、好ましい宿主細胞は、放線菌類、多くのポリケチドを豊富に産生する菌糸菌(mycelial bacteria)の階級から構築される細胞である。本発明に用いるのに特に好ましい属は、ストレプトミセスである。したがって、例えば、*S. ambofaciens*、*S. aureofaciens*、*S. avermitilis*、*S. azureus*、*S. cinnamonensis*、*S. coelicolor*、*S. curacoi*、*S. erythraeus*、*S. frad*

iae、S.galilaeus、S.glaucescens、S.hygroscopicus、S.lividans、S.parvulus、S.peucetius、S.rimosus、S.roseofulvus、S.thermotolerans、S.violaceoruber等は本発明にとって好都合な宿主を提供し、なかでもS.roseofulvusが好ましい。例えば、Hopwood, D.A. and Sherman, D.H. Ann Rev Genet (1990)24:37-66; O'Hagan, D. The Polyketide Metabolites (Ellis Horwood Limited, 1991) を、種々のポリケチド産生物および他の天然の産物の説明として参照されたい。そのような宿主細胞はまた、フレノリシン遺伝子クラスターDNA配列を単離する目的のためにも用いられ、例えば、任意の公知の寄託機関、例えば、the American Type Culture Collection (ATCC)、the Centraalbureau voor Schimmelcultures (CBS)、もしくはThe Deutsche Sammlung für Mikroorganismen und Zellkulturen GmbH (DSM) もしくはthe Journal "Industrial Property" [(1991) 1,29-40頁] に挙げられた他の寄託機関から入手することができる。例えば、ストレプトミセス・ロゼオフルバスは、ATCCからATCC 19921もしくはATCC 19805として、あるいはCBSからCBS 577.68として入手することができる。このような関係においては、本発明のDNA配列はまた、当該分野で公知の方法もしくは例えば、EP747483に記載された方法によって本明細書中に開示された配列情報に基づいて合成的に調製することもできる。

【0055】上記の細胞は、標準的な技術（例えば、相同組換え）を用いて、天然に生じる任意のPKS遺伝子をそれから欠失させることによって遺伝子操作することができる（例えば、Khosla, C. et al., 1992, Molec. Microbiol. 6:3237を参照）。集合的に置換フレノリシン遺伝子クラスターもしくは本発明のDNA配列もしくは機能的同等物をコードする遺伝子配列は、1以上の発現ベクターに、当業者に公知の方法を用いて、挿入することができる。発現ベクターは、所望のフレノリシンコード配列に作動可能に連結した制御配列を含む。本発明に用いる適切な発現系は、真核宿主細胞および原核宿主細胞内で機能する系を含む。しかし、上記に説明したように、原核生物系が好ましく、そして特にストレプトミセスと適合し得る系が特に興味深い。そのような系における使用のための制御要素は、場合によってはオペレーター配列を含むプロモーター、およびリボソーム結合部位を含む。特に有用なプロモーターは、PKS遺伝子クラスター由来の制御配列、例えば、アクチノルホジンもしくはフレノリシンクラスターの転写ユニットの5'に見いだされるようなプロモーターを含む。しかし、他の細菌プロモーター、例えば、高レベルで発現される遺伝子に由来するプロモーターもまた本発明における使用が見いだされる。例は、代謝酵素（例えば、プロテアーゼもしくはグリコシラーゼ）、生合成酵素（例えば、トリプトファンオペロン内の）、抗生物質耐性タンパク質（例えば、β-ラクタマーゼ）、もしくはバクテリオファージ

構造タンパク質をコードする遺伝子由来のプロモーター配列を含む。さらに、tacプロモーターのような天然には生じない合成プロモーター（米国特許第4,551,433号）もまた本発明の構築物における使用が見いだされる。宿主の生育に比例して、フレノリシン置換配列の発現の調節を可能にする他の調節配列もまた好ましい。調節配列は当業者には公知であり、例には調節化合物の存在を含む化学的もしくは物理的刺激に応答して切り換えられる遺伝子発現を引き起こすものを含む。他のタイプの調節要素もまたベクター内に存在してもよく、例えば、特定の生育期に発現される遺伝子から得てもよい。種々の選択マーカーを組換え発現ベクターに含むことができる。多くのマーカーは形質転換した細胞の選択に有用であることが知られており、一般に、その発現が、形質転換した細胞が適切な選択培地内で生育する場合に、その細胞に選択可能な表現型を与える遺伝子を含む。そのようなマーカーは、例えば、抗生物質耐性もしくは最少培地で生育する能力を与える遺伝子を含む。

【0056】目的の種々のフレノリシン生合成遺伝子（およびその機能的サブユニット）もしくは本発明のDNA配列は、それぞれがそれ自身のプロモーターの制御下にあるように、1以上の組換えベクターにクローン化することができる。あるいは、遺伝子は、例えば単一のプロモーターの制御下で、オペロン（フレノリシンクラスター内に天然に見いだされたものもしくは合成的に操作されたものの両方）として分類することもできる。機能的サブユニット配列は、ハイブリッドPKSが作成されるように、他のサブユニットの欠失および／または挿入を容易にする近接する制限部位を含むことができる。そのような独特な制限部位の設計は当業者に公知であり、上記の技術、例えば、部位特異的変異およびPCRを用いることにより達成することができる。本発明の組換えベクターを適切な宿主に導入する方法は、当業者に公知である。ストレプトミセスについては、プロトプラストは、細胞壁をリゾチームで消化することにより代表的に調製され、このプロトプラストを、形質転換DNAによって、その取り込みを刺激する種々の薬剤（例えば、ポリエチレングリコールおよび二価カチオン）と組合せて処理する（Hopwood DA, et al., 1985, Genetic Manipulation of Streptomyces-A Laboratory Manual, John Innes Foundation, Norwich, UK）。このプロトプラストは、実質的にもとの細胞に再生することが可能であり、形質転換体が選択される。DNAはまた、ストレプトミセスおよび他の細菌にエレクトロポレーションによって導入することができる。一旦、PKS遺伝子が発現されると、公知の技術を用いてポリケチドを産生するコロニーが同定および単離される。適切な発酵条件下で形質転換体によって産生されたポリケチドは、次に、さらにプロセスすることができる。

【0057】発酵の代表的なプロトコルは、Iwaiらによ

ってJ. Antibiotics 31:959(1978)に、およびOmura らによって米国特許第4,199,514 号に記載されている。本発明のベクターで形質転換した微生物、例えば、ストレプトミセス・ロゼオフルバス(*Streptomyces roseofulvus*)を、適切な炭素源および窒素源、例えば、炭素源については、たら肝油、コーン油、グルコース、マルトースなど、および窒素源についてはダイズ粉末、ペプトン、酵母などの存在下で培養することができる。発酵は、好氣的条件下、20~35℃で、約100~150時間培養することによって行うことができる。培養ブ罗斯は細胞と滲液に分離することができ、フレノリシンは、疎水性有機溶媒、例えば、酢酸エチル、酢酸ブチル、ベンゼンなどを用いて滲液から抽出することができる。次いで、フレノリシンを従来の技術、例えば、ゲルクロマトグラフィーにより互いに分離することができ、そして中間体化合物(エポキシおよびデオキシフレノリシン)をフレノリシンBに変換することができる。あるいは、無細胞発酵ブ罗斯(好ましくは遠心分離によって入手した)を、大気圧下、室温、pH8~10で水素化し、エポキシドをデオキシフレノリシンに変換することができる。水素化後、pHを5.5に調整し、反応混合液に60℃で空気を吹き込んでデオキシフレノリシンをフレノリシンBに酸化する。次いで、このフレノリシンBを抽出し沈降させる。

【0058】

【実施例】

#### 実施例 1

*Streptomyces*/E.coliシャトルベクターpSSVtsr およびpSSVaph の構築を図3に示す。pSG5ストレプトミセスレプリコンおよびチオストレプトン耐性遺伝子をベクターpGM160(Muth et al., 1989) からカセットとして単離した。ベクターpGM160自身はベクターpCZA168(Solenberg and Baltz, 1991)からHindIII 挿入物としてトランスポゾンを欠失させることにより入手した。ベクターpGM160をBamHIで完全に、そしてBclIで部分的に消化し、pSG5ストレプトミセスレプリコンを有する2.6kbのフラグメントをゲル単離し、Litmus28およびLitmus38(New England Biolabs) の両方に連結した。Litmusベクターのポリリンカー内でSnaBI に隣接したBamHI 部位を再構築したSG5 挿入物の方向は、プラスミドpL28SG5 およびpL38SG5 を得るように選択した。pSG5の最小レプリコンの測定に基づいて、Muth et al.(1988) によって記載されたように、必須でないDNA配列を欠失させて、pL28SG5a およびpL38SG5aを以下のように得た。

【0059】pL38SG5 をKasIで消化し、ゲル単離した4.8kbのフラグメントを再連結してpL38SG5aを得た。pL28SG5 をKasIおよびBamHI で消化し、ゲル単離した4.8kbのフラグメントを以下に示す合成リンカーに連結してpL28SG5aを得た。

GATCCGCGAGTACTGCC

GCGTCATGACGGCGCG

チオストレプトン耐性マーカーをpCZA168 から1.1kbのBclI/EcoRIフラグメントとして単離し、Litmus28のBamHI/EcoRI 部位にクローン化し、pL28tsr を得た。*S. fradiae* のネオマイシン耐性遺伝子(aph) を、1.0kbのSacII フラグメントとしてpIJ680(Hopwood et al., 1985)から単離し、pBluescript SK-(Stratagene, La Jolla, CA, USA)にクローン化してpBSaphを得た。

【0060】pL28tsr をNdeIで切断し、この部位をE.coliのDNAポリメラーゼIのKlenowフラグメントで満たすことによって平滑化した。次いで、このベクターをHindIII で切断し、得られた1.1kbのフラグメントをゲル単離した。pL28SG5aをHindIII およびHpaIで切断し、5.2kbのフラグメントをゲル単離した。これら2つのフラグメントを、次に、連結してpSSVtsr\*を得た。pSSVtsr を、pSSVtsr\*から、BamHI およびBglII を用いて切断し再連結することにより誘導した。

【0061】pBSaphをSpeIで切断し、この部位をE.coliのDNAポリメラーゼIのKlenowフラグメントで満たすことによって平滑化した。このベクターを、次に、Tsp509I切断し、1.1kbのフラグメントをゲル単離した。pL38SG5aをEcoRI およびHpaIで切断し、5.2kbのフラグメントをゲル単離した。これら2つのフラグメントを、次いで、連結してpSSVaph を得た。

#### 【0062】実施例 2

ermE\* およびtipAプロモーターを便利な制限部位で合成した。一連のオーバーラップしたオリゴヌクレオチドをABI 392 DNA 合成機で合成し、続いてT4ポリヌクレオチドキナーゼでリン酸化した。次いで、これらのオリゴヌクレオチドを90℃にまで加熱し、ゆっくりと冷却することによりアニーリングした。アニーリング後、このフラグメントをT4リガーゼを用いて室温で4時間連結し、その後KpnIおよびBamHI で切断した。ゲル単離したフラグメントをLitmus29のKpnIおよびBamHI 部位にクローン化し、ベクターpTipA およびpErmE を得た(図5)。

【0063】フレノリシンクラスター内の種々の遺伝子をPCRによって増幅し、pTipA および/またはpErmE にクローン化し、次いで、以下の遺伝子G、すなわち転写アクチベーターについて詳細に記載するのと実質的に同じ手法でシャトルベクターに移し入れた(図5)。選択した遺伝子の過剰発現用のベクターは、したがって、次に記載するようにして構築した。遺伝子Gを以下に示すプライマーを用いてPCRにより増幅した。得られたPCR産物を、BamHI (およびその産物をpTipA に連結する場合にはそれに加えてNdeI)を用いて消化し、ゲル単離し、pErmEBamHI 部位もしくはpTipA のBamHI 部位とNdeI部位との間にクローン化し、それぞれpErmTAおよびpTipTAを得た。

ermE\* に融合するための遺伝子Gを増幅するための順序

向(5')PCRプライマー:CGGGATCCAGCGGTGGAGATCAAGTACATGGGTCAGTTGACC

tipAに融合するための遺伝子Gを増幅するための順方向(5')PCRプライマー:CGCATATGGAGATCAAGTACATGGTTCAGTTGACC

pErmEもしくはpTipAにクローン化するための遺伝子Gを増幅するための逆方向(3')PCRプライマー:GCGGATCCGTGTCTAGTCTGTCGAGCGCGCCGCGGTGGC

【0064】pErmTAおよびpTipAベクターをAseIおよびStuIで切断し、大きな方のフラグメント(2.8kb)をゲル単離した。同時に、pSSVtsrおよびpSSVaphベクターをAseIおよびSnaBIで切断し、そして大きな方のフラグメント(3.9kb)をゲル単離した。pErmTAおよびpSSVtsrフラグメントと一緒に連結してSSVtsr-ermTAを得た。pTipAおよびpSSVaphフラグメントと一緒に連結してSSVaph-tipAを得た。tipAプロモーターによって遺伝子を発現する構築物のために、ネオマイシン耐性遺伝子を選択用に用い、その結果、チオストレプトンを誘導用に用いることができる。AseI、SnaBI、StuI、BamHIおよびNdeI部位はフレノリシンクラスター内にないので、任意のフレノリシン遺伝子をこの方法を用いて発現させることができる。

【0065】クラスター内の特定の遺伝子を破壊するために設計されたベクターを調製するために、これらの遺伝子に内在するフラグメントをLitmus28もしくは38にクローン化し、そして同じ方法を用いてフラグメントをシャトルベクターに移し入れた。例えば、ほとんどの遺伝子Lから伸びる1.6kbのSalIフラグメントと遺伝子Mの5'末端をLitmus38に連結し、得られたベクターをpSSVtsrにつないでSSVtsr-デルタLMを得た。

#### 【0066】実施例 3

ストレプトミセスを形質転換するために用いられるベクターをメチル化欠損(dam-,dcm-)E.coli株内で増殖させた。この株から単離したベクターDNAを用いてS.lividans、S.coelicolor、S.roseofulvusを形質転換した。

S.roseofulvus含むストレプトミセスは、先に記載されたもの(Hopwood D, et al., 1985, Genetic Manipulation of Streptomyces-A Laboratory Manual, The John Innes Foundation, Norwich, UK)を改変した手法を用いて形質転換した。改変YEME培地25ml(1L当たり:3gのDifco イーストエキストラクト、5gのDifco バクトーペプトン、3gのDifco モルトエキストラクト、10gのグルコース、250gのスクロース;オートクレーブ後にMgCl<sub>2</sub>およびグリシンをそれぞれ終濃度5mMおよび0.1%で添加)を250ml容のバッフルフラスコにいれ、それに孢子懸濁物を接種し、22ORPMで振盪しながら5~8日間30℃でインキュベーションした。得られた菌糸を10.3%のスクロースで2回洗浄し、0.5mg/mlのリゾチームおよび0.5mg/mlのアクロモペプチダーゼ(achromopeptidase)を含む4mlの溶菌緩衝液(Hopwood et al., 1985, 前出)に再懸濁した。細胞壁を10~15分間室温で消化し、顕微鏡でモニターした。残りの菌糸フラグメントを、消化物を綿栓を通してさせることにより取り除き、プロトプラストを新鮮なプロトプラスト緩衝液(Hopwood et al., 1985, 前出)で2回洗浄し、-80℃で保存した。

【0067】形質転換は解凍したプロトプラストの50μlのアリコート、0.5~5μgのプラスミドDNAと一緒に、10μlの微量元素と形質転換緩衝液中25%のPEG 6000を500μl含む溶液(Hopwood et al., 1985, 前出)中で、室温で1分間インキュベーションすることにより行った。形質転換体を500μlのプロトプラスト緩衝液と混合し、そして100~500μlのアリコートを再生プレート(10.3%のスクロースを含むCM1-2)上に広げた。30℃で24時間インキュベーションした後、チオストレプトン(終濃度30mg/Lを得る)もしくはネオマイシン(終濃度10mg/Lを得る)を含む軟寒天培地にプレートかぶせ、形質転換した細胞を単離した。

#### 【0068】

表1

遺伝子(配列番号)	コードされるタンパク質の性質
A(SID:1)	80kD、強い相同性を持たない非膜性タンパク質
B(SID:2)	Aに移行的にカップリングしている膜性タンパク質; 輸送に関与し得る
C(SID:3)	6個の膜貫通ドメインを有するタンパク質
D(SID:4)	Cに移行的にカップリングしている
E(SID:5)	ABCトランスポーターの可溶性ATP-結合成分 機能は不明; データベース内に相同性なし; 膜性タンパク質ではない
F(SID:6)	actVA-ORF1に関連した推定放出ポンプ; 12-14個の膜貫通ドメイン
G(SID:7)	actII-ORF4に関連した転写アクチベーター
H(SID:8)	遺伝子Iに移行的にカップリングしている

I (SID:9)	fabHのホモログ; 4炭素スターターユニットを合成し得る
J (SID:10)	上記スターターユニットを運ぶためにおそらく 上記fabHホモログによって用いられるACP
K (SID:11)	アシルトランスフェラーゼ; ACP(J) からPKSに スターターユニットを輸送し得る
L (SID:12)	PKSケトアシルシンターゼサブユニット
M (SID:13)	別のPKSサブユニット; いわゆる「鎖長因子」
N (SID:14)	PKSに用いられるACP
O (SID:15)	actVI-ORF5に関連; おそらくヘミケタールデヒドラーゼ
P (SID:16)	actIIIに関連するケトレダクターゼ
Q (SID:17)	actVIIに関連するシクラーゼ/デヒドラーゼ
R (SID:18)	actIV に関連するシクラーゼ/デヒドラーゼ
S (SID:19)	特定のactVI 遺伝子に弱く関連するオキシドレダクターゼ
T (SID:20)	actVA-ORF5に関連するキノン-形成ヒドロキシラーゼ

## 【0069】

## 【配列表】

配列番号: 1:

配列の特性:

配列の長さ: 770アミノ酸

配列の型: アミノ酸

鎖の数: 一本鎖

トポロジー: 直鎖状

配列の種類: タンパク質

ハイボセティカル: NO

アノイズ: NO

## 配列

```

Met Ala Gly Ser Gly Tyr Ala Tyr Tyr Gln His Leu Ser Gly Asn Ile
1           5           10           15
Asp Lys Ile Asp Val Gly Asp Ala Gly Asn Lys Asp Ala Ala Pro Asp
          20          25          30
Gly Pro Ile Asn Ile Leu Ile Ile Gly Thr Asp Lys Arg Thr Gly Lys
          35          40          45
Gly Asn Glu Gly Cys Gly Gly Lys Asp Ser Pro Gly His Ala Asp Thr
          50          55          60
Asn Ile Leu Leu Arg Val Ser Ala Asp Arg Thr Asn Thr Thr Gly Leu
65          70          75          80
Ser Ile Pro Arg Asp Leu Ile Thr Asn Ile Pro Asp Cys Leu Thr Thr
          85          90          95
Gln Asp Asp Gly Ser Lys Lys Thr Ile Pro Gly Thr Gln Asn Val Arg
          100         105         110
Phe Asn Thr Ser Leu Gly Gln Glu Gly Arg Asp Pro Gly Cys Thr Met
          115         120         125
Arg Thr Val Thr Glu Leu Thr Gly Leu Lys Val Asp His Phe Met Met
          130         135         140
Ala Asp Phe Asn Ala Val Lys Asn Leu Thr Thr Ala Val Asn Gly Val
145         150         155         160
Glu Val Cys Val Ala Lys Asp Val Asp Asp Pro Asp Ser His Leu Lys
          165         170         175
Leu Ser Ala Gly Thr His Lys Val Gln Gly Glu Gln Ala Leu Ala Phe
          180         185         190
Val Arg Thr Arg His Ser Phe Gly Asn Gln Gly Asp Leu Asp Arg Ile
          195         200         205
Lys Val Gln Gln Gln Phe Leu Gly Ser Leu Ala Arg Gln Leu Lys Ser
210         215         220
Glu Asp Thr Leu Thr Ser Pro Lys Lys Leu Tyr Lys Val Ala Glu Ala

```



225                      230                      235                      240  
 Ala Thr Asp Ala Leu Thr Val Asp Ser Gly Ile Gly Ser Ile Thr Lys  
                                  245                      250                      255  
 Leu Met Ser Leu Ala Lys Glu Leu Gln His Ile Asn Pro Lys Asn Ile  
                                  260                      265                      270  
 Thr Phe Val Thr Leu Pro Val Val Asp Asn Pro Ala Glu Lys Val Lys  
                                  275                      280                      285  
 Ala Thr Val Val Leu Asn Glu Thr Asp Ala Asp Pro Gln Gln Ser Ala  
                                  290                      295                      300  
 Leu Gly Gln Ser Leu Asp Val Gly Arg Gln Leu Val Asp Ser Leu Thr  
 305                      310                      315                      320  
 Asp Gln Asp Pro Arg Asp Gly Lys Thr Val Pro Trp Leu Ala Thr Arg  
                                  325                      330                      335  
 Trp Lys Ala Asp Pro Glu Ala Thr Arg Phe Thr Phe Thr Leu Arg Ala  
                                  340                      345                      350  
 Gly Ala Thr Phe Ser Asp Gly Thr Pro Val Asp Ala Arg Ala Val Lys  
                                  355                      360                      365  
 Ala Asn Phe Asp Ala Val His Glu Leu Gly Ala Ala Ala Ser Arg Gly  
 370                      375                      380  
 Ala Val Tyr Leu Asp Gly Tyr Arg Glu Thr Arg Val Ser Gly Ala Arg  
 385                      390                      395                      400  
 Thr Leu Thr Val Val Phe Asp Lys Pro Asn Ala Gln Phe Leu Arg Gly  
                                  405                      410                      415  
 Thr Ser Thr Val Ser Leu Gly Leu Leu Ser Pro Gly Ser Leu Arg Arg  
                                  420                      425                      430  
 Thr Pro Gln Glu Arg Cys Thr Gly Arg Leu Val Gly Ser Gly Pro Phe  
                                  435                      440                      445  
 Val Leu Asp Arg Tyr Arg Pro Asn Thr Ser Val Thr Leu Asp Arg Arg  
                                  450                      455                      460  
 Lys Gly Tyr Ser Trp Gly Ser Arg Leu Trp Gln Arg Glu Gly Gly Ala  
 465                      470                      475                      480  
 Tyr Leu Glu Gly Val Glu Tyr Arg Ile Val Pro Glu Asn Thr Thr Arg  
                                  485                      490                      495  
 Ser Gly Ala Leu Ser Ala Gly Gln Leu Asp Val Ala Thr Ala Leu Ala  
                                  500                      505                      510  
 Pro Gln Asp Arg Glu Arg Phe Ser Ala Pro Gly Trp Ser Leu Leu Thr  
                                  515                      520                      525  
 Arg Thr Ala Pro Gly Val Asp Leu Ser Leu Tyr Val Asn Ala Arg Arg  
                                  530                      535                      540  
 Thr Ala Leu Arg Glu Ala Ala Val Arg Gln Ala Leu Gln Lys Gly Ile  
 545                      550                      555                      560  
 Asp Arg Glu Ala Val Ala Thr Thr Phe Leu Ser Ser Arg Lys Leu Ala  
                                  565                      570                      575  
 Ala Thr Ser Val Leu Ser Ser Thr Thr Pro Gly Tyr Thr Asp Leu Gly  
                                  580                      585                      590  
 Asp Arg Leu Ala His Asp Pro Ala Gly Ala Arg Arg Leu Leu Asp Lys  
                                  595                      600                      605  
 Ala Gly Trp Arg Pro Gly Ala Asp Gly Ile Arg Val Lys Asn Gly Val  
 610                      615                      620  
 Arg Leu Arg Leu Asp Ala Val Phe Val Arg Gln Gln Ser Leu Glu Leu

```

625          630          635          640
Val Gln Gln Gln Leu Lys Asp Ile Gly Val Glu Leu Arg Leu Arg Gln
          645          650          655
Leu Thr Val Ser Arg Phe Pro Glu Val Leu Ala Ala Gly Ser Tyr Asp
          660          665          670
Leu Ser Leu Gln Ser Ala Asn Arg Ala Asp Pro Asp Val Leu Thr Thr
          675          680          685
Ala Phe Ala Gly Gly Thr Pro Val Ala Asp Ala Arg Leu Arg Ser Glu
          690          695          700
Leu Arg Arg Ala Thr Ser Ser Thr Asp Glu Ala Thr Arg Ser Ser Leu
705          710          715          720
Phe Ala Ala Ala Gln Arg Arg Leu Val Asp Glu Gly His Val Leu Pro
          725          730          735
Leu Asn Glu Thr Glu Glu Thr Ala Ala Leu Ser Thr Arg Val His Gly
          740          745          750
Leu Thr Arg Asp Ala Ser Asn Arg Leu Val Leu His Asp Thr Trp Thr
          755          760          765
Thr Gly
          770

```

【 0 0 7 0 】配列番号 : 2 :

配列の特性 :

配列の長さ : 3 2 3 アミノ酸

配列の型 : アミノ酸

鎖の数 : 一本鎖

トポロジー : 直鎖状

配列の種類 : タンパク質

ハイポセティカル : NO

アンチセンス : NO

## 配列

```

Met Thr Ala Arg Tyr Leu Ala Arg Arg Leu Gly Arg Val Val Leu Val
1          5          10          15
Val Trp Ala Ala Tyr Thr Leu Ser Phe Ala Val Leu Tyr Leu Leu Pro
          20          25          30
Gly Asp Pro Val Gln Thr Met Leu Ser Gly Ala Ala Gly Gly Asp Gly
          35          40          45
Ala Ala Val Asp Pro His Glu Ala Gln Arg Leu Arg His Thr Leu Gly
          50          55          60
Leu Asp Arg Pro Leu Ala Val Gln Tyr Thr Ser Met Leu Gly His Ala
65          70          75          80
Leu Arg Gly Asp Leu Gly Thr Ser Ile Arg Ser Gly Ala Pro Val Arg
          85          90          95
Gly Gln Leu Ala Gln Ala Leu Pro Asp Thr Leu Ser Val Ala Leu Pro
          100          105          110
Ala Leu Val Leu Ser Val Leu Val Ala Leu Cys Leu Ala Leu Leu Gly
          115          120          125
Ala Trp Pro Arg Arg Arg Ala Leu Arg Arg Ala Ala Thr Ala Leu Pro
          130          135          140
Ser Leu Gly Thr Ala Met Pro Ser Phe Trp Leu Gly Leu Leu Leu Ala
145          150          155          160
Gln Trp Val Ser Phe Arg Trp Gly Leu Leu Pro Ala Thr Gly Gly Gly
          165          170          175
Arg Ser Pro Arg Ala Thr Leu Leu Ala Ala Leu Thr Leu Ala Leu Pro
          180          185          190
Ile Gly Cys Val Leu Ala Gln Val Leu Gly Arg Gly Leu Arg Ala Ala

```

```

          195          200          205
Leu Ala Glu Pro Tyr Ala Asp Val Ala Arg Ser Arg Gly Ala Gly Arg
    210          215          220
Ala Arg Leu Leu Leu Ala Arg Ala Leu Arg Asn Ala Ser Val Ala Ala
    225          230          235          240
Leu Ala Leu Leu Gly Val Val Cys Gly Gln Leu Leu Ala Gly Ala Val
          245          250          255
Leu Val Glu Thr Val Phe Ala Arg Gly Gly Ile Gly Arg Leu Ala Met
          260          265          270
Asp Ala Val Thr Tyr Gln Asp Leu Pro Val Val Gln Gly Val Val Val
    275          280          285
Leu Ala Ala Leu Val Ala Ala Leu Val Asn Leu Val Val Asp Leu Leu
    290          295          300
Leu Pro Leu Leu Glu Pro Arg Thr Ala Ser Glu Ala Ala Asp Ala Val
    305          310          315          320
Pro Ala His

```

【 0 0 7 1 】配列番号 : 3 :

配列の特性 :

配列の長さ : 2 7 4 アミノ酸

配列の型 : アミノ酸

鎖の数 : 一本鎖

トポロジー : 直鎖状

配列の種類 : タンパク質

ハイポセティカル : NO

アンチセンス : NO

## 配列

```

Met Ala Leu Arg Arg Val Ala Ala Leu Trp Arg Ala Pro Gly Leu Ala
1          5          10          15
Leu Ser Leu Leu Val Leu Leu Leu Val Leu Gly Trp Ala Leu Leu Pro
    20          25          30
Gly Leu Phe Thr Ala Ala Asp Pro Leu Arg Ala Asp Ala Ala His Arg
    35          40          45
Leu Leu Ala Pro Gly Ala Gly His Pro Leu Gly Ala Asp His Val Gly
    50          55          60
Arg Asp Leu Tyr Ala Arg Val Val His Gly Thr Ala Arg Ser Leu Gly
    65          70          75          80
Thr Ala Phe Ala Ala Val Ala Leu Gly Val Leu Ala Gly Gly Ala Leu
          85          90          95
Gly Ala Val Ala Gly Val Ala Gly Arg Ala Val Asp Ala Val Val Met
          100          105          110
Arg Val Val Asp Val Leu Leu Ala Val Pro Gly Leu Leu Leu Ser Leu
    115          120          125
Ala Val Val Ser Ala Leu Gly Phe Gly Thr Ala Gln Val Ala Cys Ala
    130          135          140
Val Gly Val Gly Thr Val Gly Gly Ile Ala Arg Val Ser Arg Ala Gln
    145          150          155          160
Val Arg Arg Val Arg Gly Gly Glu Tyr Val Glu Ala Ala Arg Leu Ala
          165          170          175
Gly Val Ala Gly Pro Leu Ile Leu Leu Arg His Ile Val Pro Asn Ala
          180          185          190
Ala Pro Pro Val Leu Ala Leu Ala Val Thr Glu Cys Gly Thr Ala Val
          195          200          205
Leu Gly Val Ala Ser Leu Gly Phe Leu Gly Phe Gly Ala Pro Pro Pro
    210          215          220

```

Ala Pro Glu Trp Gly Ala Leu Ile Ser Thr Gly Arg Asp Tyr Leu Val  
 225 230 235 240  
 Ser Ala Trp Trp Leu Thr Thr Leu Pro Gly Leu Val Leu Val Ala Leu  
 245 250 255  
 Val Val Ala Leu His Arg Val Gly Arg Ala Leu Glu Arg Glu Glu Arg  
 260 265 270  
 Thr Gly

【0072】配列番号：4：

トポロジー：直鎖状

配列の特性：

配列の種類：タンパク質

配列の長さ：524アミノ酸

ハイボセティカル：NO

配列の型：アミノ酸

アンチセンス：NO

鎖の数：一本鎖

## 配列

Met Thr Pro Ala Asp Lys Pro Thr Asp Glu Arg Ser Pro Val Leu Asp  
 1 5 10 15  
 Leu Ser Gly Val Ser Val Ala Tyr Gly Thr Arg Thr Val Leu His Gly  
 20 25 30  
 Ile Asp Leu Arg Leu Ala Pro Gly Gln Val Leu Ala Val Leu Gly Ala  
 35 40 45  
 Ser Gly Ser Gly Lys Ser Thr Leu Ala Gln Ala Ala Leu Gly Leu Leu  
 50 55 60  
 Pro Pro Gly Gly Arg Val Thr Ala Gly Arg Val Thr Val Ala Gly His  
 65 70 75 80  
 Asp Ile Thr Ala Leu Ala Pro His Arg Leu Arg Ala Leu Arg Gly Thr  
 85 90 95  
 Val Thr Gly Leu Val Pro Gln Asp Gln Ala Val Ser Leu Asp Pro Leu  
 100 105 110  
 Val Arg Val Gly Ala Gln Val Thr Glu Thr Leu Arg Ala His Arg Leu  
 115 120 125  
 Glu Asp Arg Arg Glu Ala Ala Arg Arg Ala Val Pro Leu Leu Gly Glu  
 130 135 140  
 Ala Gly Ile Glu Ala Pro Gly Pro Leu Ala Arg Ala Tyr Pro His Ala  
 145 150 155 160  
 Leu Ser Gly Gly Gln Arg Gln Arg Val Leu Val Ala Gly Ala Phe Ala  
 165 170 175  
 Ala Arg Pro Pro Leu Val Val Ala Asp Glu Pro Thr Ser Ala Leu Asp  
 180 185 190  
 Ala Thr Val Arg Arg Arg Val Met Asp Arg Phe Ala Ala Leu Val Ala  
 195 200 205  
 Ala His Gly Thr Ala Val Leu Leu Val Thr His Asp Phe Arg Leu Ala  
 210 215 220  
 Arg Glu Arg Ala Asp Gln Val Ala Val Leu Ala Asp Gly Arg Leu Val  
 225 230 235 240  
 Glu Ser Gly Pro Ala Ala Arg Val Leu Asp Arg Pro Ala His Pro Tyr  
 245 250 255  
 Thr Arg Arg Leu Thr Gly Ala Gly Arg Arg Val Ala Ala Arg Gly Thr  
 260 265 270  
 Ala Pro Arg Ala Ser Gly Thr Pro Val Val Arg Ala Arg Asp Leu Val  
 275 280 285  
 Lys Glu Tyr Arg Arg Asp Gly Arg Arg Val Arg Ala Val Asp Gly Val

```

      290              295              300
Gly Phe Thr Val Arg Glu Gly Glu Phe Phe Ala Leu Val Gly Glu Ser
305              310              315              320
Gly Ser Gly Lys Ser Thr Thr Ala Arg Leu Val Thr Gly Leu Thr Ala
      325              330              335
Pro Thr Ser Gly Ala Val Glu His Ala Pro Ala Pro Val Arg Pro Gln
      340              345              350
Leu Val Gln Gln Ser Pro Tyr Ala Ala Phe Asp Pro Arg Trp Thr Val
      355              360              365
Arg Arg Ile Val Glu Glu Pro Leu Arg Ala Arg His Val Pro Gly Ala
      370              375              380
Arg Arg Arg Ala Arg Leu Arg Glu Leu Leu Ala Leu Val Gly Leu Asp
385              390              395              400
Glu Glu Leu Leu Ala Arg Arg Pro Arg Glu Leu Ser Gly Gly Gln Arg
      405              410              415
Gln Arg Val Ala Leu Ala Arg Ala Leu Ala Pro Glu Pro Arg Leu Leu
      420              425              430
Val Cys Gly Glu Pro Val Ala Ala Leu Asp Pro Val Ala Arg Glu Arg
      435              440              445
Val Val His Leu Leu Glu Arg Leu Arg Ala Glu Leu Gly Leu Thr Cys
      450              455              460
Leu Phe Val Ser His Glu Leu Asp Val Val Arg Arg Leu Cys Gly Arg
465              470              475              480
Val Ala Val Met Arg Gly Gly Arg Leu Leu Glu Ser Gly Pro Val Gly
      485              490              495
Glu Val Leu Ser Ala Pro Ser Gln Pro Tyr Thr Arg Ala Leu Leu Ala
      500              505              510
Ala Glu Ala Gly Pro Ser Asp Thr Pro Gly Ala Gly
      515              520

```

【 0 0 7 3 】配列番号 : 5 :

配列の特性 :

配列の長さ : 2 1 6 アミノ酸

配列の型 : アミノ酸

鎖の数 : 一本鎖

トポロジー : 直鎖状

配列の種類 : タンパク質

ハイボセティカル : NO

アンチセンス : NO

## 配列

```

Met Asn Glu Ile Thr Val Glu Ile Trp Thr Asp Val Val Cys Pro Trp
1              5              10              15
Cys Tyr Ile Gly Lys Arg Arg Phe Glu Arg Ala Leu Ala Ala Phe Asp
      20              25              30
Ala Lys Glu Asp Val Arg Val His Trp Arg Ser Phe Glu Leu Asp Pro
      35              40              45
Ala Ala Leu Arg Val Thr Asp Glu Thr Ile Pro Glu Arg Met Leu Arg
      50              55              60
Arg Gln Gly Ile Pro Pro Glu Gln Ala Ala Glu Leu Leu Ala Gly Val
65              70              75              80
Ser Ala Gln Ala Glu Ala Glu Gly Leu Glu Tyr His Leu Asp Arg Ala
      85              90              95
Arg Pro Cys Asn Thr Phe Asp Ala His Arg Leu Ala His His Ala Gly
      100              105              110
Thr Arg Gly Leu Ala Glu Thr Phe Gln Glu Arg Leu Met Cys Ala Tyr

```

115	120	125
Thr Ala Glu Gly Val Ser	Val Gly Asp His Pro	Thr Leu Leu Ala Leu
130	135	140
Ala Glu Glu Ala Gly Leu	Asp Ala Ala Ala Ala	Glu Val Leu Ala
145	150	155
Gly Asp Ala His Ala Glu	Asp Val Arg Ala Asp	Glu Asp Arg Ala Ala
165	170	175
Arg Leu Gly Val Gly Gly	Val Pro Ala Phe Val	Ile Gly Gly Arg Trp
180	185	190
Ser Val Ser Gly Ala Gln	Pro Ala Glu Leu Leu	Thr Gly Leu Leu Glu
195	200	205
Arg Ala Arg Thr Ala Ala	Ala Ala	
210	215	

【0074】配列番号：6：

配列の特性：

配列の長さ：517アミノ酸

配列の型：アミノ酸

鎖の数：一本鎖

トポロジー：直鎖状

配列の種類：タンパク質

ハイボセティカル：NO

アンチセンス：NO

## 配列

Met	Ser	Ser	Ser	Pro	Pro	Ala	Pro	Ala	Thr	Pro	Gly	Val	Ala	Pro	His
1				5					10					15	
Ser	Pro	Pro	Ala	Pro	Arg	Leu	Gly	Leu	Val	Leu	Leu	Val	Cys	Cys	Leu
			20					25					30		
Ala	Gln	Phe	Leu	Val	Thr	Leu	Ser	Val	Ala	Ile	Val	Asn	Val	Ala	Leu
		35					40				45				
Pro	Asp	Ile	Gln	Arg	Gly	Leu	Gly	Phe	Ser	Ala	Glu	Ser	Leu	Gln	Trp
	50				55				60						
Val	Val	Asn	Ala	Tyr	Thr	Val	Thr	Phe	Ala	Gly	Phe	Leu	Leu	Leu	Gly
65				70				75				80			
Gly	Arg	Ile	Ala	Asp	Leu	Phe	Gly	Arg	Arg	Arg	Ile	Phe	Leu	Ala	Gly
			85					90				95			
Val	Ala	Leu	Phe	Ala	Leu	Ala	Ser	Leu	Ala	Gly	Gly	Leu	Ser	Gln	Asn
		100					105					110			
Ala	Gly	Thr	Leu	Val	Ala	Ala	Arg	Ala	Val	Gln	Gly	Leu	Ala	Ala	Ala
	115						120				125				
Val	Ile	Ala	Pro	Thr	Thr	Leu	Ala	Val	Leu	Gly	Thr	Ser	Phe	Lys	Asp
	130						135				140				
Pro	His	Gln	Arg	His	Arg	Ala	Phe	Gly	Ala	Trp	Gly	Ala	Val	Ser	Gly
145				150				155				160			
Ala	Gly	Gly	Ala	Phe	Gly	Ala	Leu	Ala	Gly	Gly	Ala	Leu	Thr	Asp	Ala
			165					170				175			
Phe	Ser	Trp	Arg	Trp	Val	Leu	Phe	Val	Asn	Leu	Pro	Ile	Gly	Val	Leu
	180						185					190			
Leu	Leu	Ala	Gly	Ile	Ala	Trp	Gly	Ile	Ser	Glu	Leu	Arg	His	Ala	Gly
	195						200				205				
Glu	Asp	Arg	Arg	Ile	Asp	Val	Ala	Gly	Ala	Leu	Thr	Val	Thr	Leu	Gly
	210						215				220				
Leu	Leu	Ala	Leu	Val	Leu	Gly	Ile	Val	Gln	Ser	Gly	Pro	His	Gly	Trp
225				230				235				240			
Gly	Ser	Ala	Ala	Thr	Leu	Val	Pro	Leu	Leu	Gly	Gly	Leu	Ala	Leu	Leu

```

      245              250              255
Gly Ala Phe Val Leu Val Glu Gly Arg Phe Ala Pro Gln Pro Leu Ile
      260              265              270
Pro Leu Gly Ile Phe Arg Ser Arg Ser Val Val Ala Ala Asn Val Val
      275              280              285
Ala Met Thr Ser Gly Ala Ala Leu Phe Ser Met Phe Tyr Phe Leu Thr
      290              295              300
Leu Phe Leu Asn Gln Val Arg Asp Tyr Ser Pro Leu Arg Thr Gly Phe
305              310              315              320
Ala Tyr Leu Pro Leu Ala Leu Ala Ile Met Val Ala Ala Gln Phe Ser
      325              330              335
Ala Ala Leu Val Arg Val Leu Gly Pro Arg Thr Thr Leu Leu Val Ser
      340              345              350
Met Ala Leu Thr Ala Ala Gly Leu Leu Trp Leu Ser Arg Leu Thr Glu
      355              360              365
Asp Ser Gly Phe Ala Gly Gly Leu Leu Gly Pro Thr Leu Val Val Gly
      370              375              380
Ile Gly Gln Gly Ile Ser Met Ser Ala Ser Ala Ile Ala Gly Val Ala
385              390              395              400
Gly Val Arg Pro Gln Gln Ala Gly Leu Ala Ser Gly Leu Leu Asn Ala
      405              410              415
Thr Arg Gln Leu Gly Gly Ala Leu Gly Leu Ala Val Val Ala Ala Val
      420              425              430
Ala Thr Ser Arg Ala Asp Gly Leu Leu Asp Gly Val Ala Arg Pro Thr
      435              440              445
Ala Glu Leu Ala Arg His Ala Gln Ala Ser Gly His Pro Leu Ser Ile
      450              455              460
Ala Val Ala Ala Ala Leu Ser Ala Val Gly Leu Leu Ala Ser Leu Ala
465              470              475              480
Ala Pro Gly Arg Ser Pro Ala Pro Thr Gly Thr Arg Thr Gly Gly Asp
      485              490              495
Ser Ala Ala Pro Ala Pro Ala Ala Ala Pro Ala Ala Thr Gly Thr Thr
      500              505              510
Gly Pro Gly Glu Ile
      515

```

【 0 0 7 5 】配列番号 : 7 :

配列の特性 :

配列の長さ : 2 8 2 アミノ酸

配列の型 : アミノ酸

鎖の数 : 一本鎖

トポロジー : 直鎖状

配列の種類 : タンパク質

ハイポセティカル : NO

アンチセンス : NO

## 配列

```

Met Glu Ile Lys Tyr Met Gly Gln Leu Thr Met Arg Trp Glu Gly Arg
1              5              10              15
Glu Lys Leu Pro Ser Ala Arg Lys Pro Arg Thr Val Leu Ala Leu Leu
      20              25              30
Leu Leu Asn Asp Lys Thr Pro Val Thr Thr Ser Ala Leu Ile Thr Glu
      35              40              45
Leu Trp Gly Glu Asn Pro Pro Arg Ser Ala Leu Thr Thr Leu Gln Thr
      50              55              60
Tyr Ile Leu Gln Leu Arg Lys Cys Leu Ala Ala Met Ser Gly Arg Ser

```

```

65          70          75          80
Leu Ala Cys Ile Ser Glu Lys Thr Leu Val Thr Trp Pro Cys Gly Tyr
          85          90          95
Leu Ala Arg Leu Pro Ala Asp Ala Thr Ser Asp Val Ala Glu Phe Arg
          100         105         110
Arg Phe Ala Arg Glu Gly Arg Glu Ala Glu Arg Arg Gly His Leu Ala
          115         120         125
Glu Ala Val Arg Ser Tyr Arg Ala Ala Leu Ser Leu Ser Gln Gly Pro
          130         135         140
Leu Leu Ala Asp Ile Glu His Gly Pro Leu Leu Arg Ala Glu Ala Val
          145         150         155         160
Arg Met Glu Glu Cys Arg Leu Ser Leu Val Glu Arg Ser Ile Glu Gly
          165         170         175
Asp Leu Leu Leu Gly Arg His Arg Glu Val Val Ser Glu Leu Ser Ala
          180         185         190
Leu Val Ala Gln Tyr Pro Tyr His Glu Gln Leu Thr Gly Gln Leu Met
          195         200         205
Val Ala Leu Val Arg Cys Gly Arg Arg Gln Asp Ala Leu Ala Val His
          210         215         220
Gln Arg Leu Arg Ala Arg Met Val Glu Asp Leu Gly Leu Glu Pro Ser
          225         230         235         240
Ser His Leu Arg Ala Leu Gln Ser Ala Val Leu Ser Gly Glu Pro Leu
          245         250         255
Pro Gly Pro Pro Gly Thr Gly Gly Glu Ile Pro Thr Pro Tyr Ala Gly
          260         265         270
Ala Phe Ala Thr Ala Ala Arg Ser His Asp
          275         280

```

【 0 0 7 6 】配列番号 : 8 :

配列の特性 :

配列の長さ : 1 7 1 アミノ酸

配列の型 : アミノ酸

鎖の数 : 一本鎖

トポロジー : 直鎖状

配列の種類 : タンパク質

ハイボセティカル : NO

アンチセンス : NO

## 配列

```

Met Asn Glu Thr Arg Thr Ala Ala Arg Thr Gly Gln Val Gly Pro Val
1          5          10          15
Asp Ala Glu Gly Phe Arg Ala Ala Met Ser Cys Phe Pro Ala Gly Val
          20          25          30
Val Leu Val Thr Thr Arg Glu Glu Asp Gly Thr Pro Arg Gly Phe Thr
          35          40          45
Ala Ser Ser Phe Cys Ser Val Ser Leu Ala Pro Pro Leu Val Ser Val
          50          55          60
Cys Gln Gly Thr Gly Ala Gln Ser Tyr Gly Ala Phe Gln Glu Cys Ala
65          70          75          80
Arg Phe Ala Val Ser Val Leu Arg Ser Gly Gln Arg Glu Leu Ala Ser
          85          90          95
Arg Phe Ala Thr Arg Gly Ala Asp Lys Phe Gly Gly Gly Gly Leu Val
          100         105         110
Ala Leu Glu Gly Ser Gly Leu Leu Val Ala Ala Asp Ala Leu Val Thr
          115         120         125
Leu Glu Cys Ala Val His Ala Arg His Leu Ala Gly Asp His Val Ile

```



130                      135                      140  
 Leu Val Gly Glu Val Arg Gly Val Gly Gln Gly Glu Gly Glu Pro Leu  
 145                      150                      155                      160  
 Val His Trp Glu Arg Gly Phe Arg Ala Leu Arg  
                                  165                      170

【0077】配列番号：9：

配列の特性：

配列の長さ：336アミノ酸

配列の型：アミノ酸

鎖の数：一本鎖

トポロジー：直鎖状

配列の種類：タンパク質

ハイポセティカル：NO

アンチセンス：NO

配列

Met Thr Gly Glu Ala Glu Met Leu Gly Thr Arg Pro Val Val His Ser  
 1                      5                      10                      15  
 Arg Leu Leu Gly Val Gly Gly Tyr Arg Pro Arg Arg Ser Val Asp Asn  
                                  20                      25                      30  
 Ala Glu Leu Cys Ala Thr Val Ala Ser Thr Pro Glu Trp Ile Glu Thr  
                                  35                      40                      45  
 Arg Ser Gly Ile Arg Ala Arg Gly Phe Ala Ala Pro Asp Glu Thr Leu  
                                  50                      55                      60  
 Arg Phe Met Gly Arg Ala Ala Ala Glu Lys Ala Leu Ala Arg Ala Gly  
 65                      70                      75                      80  
 Val Leu Pro Asp Gly Ile Asp Leu Val Leu Val Ala Ser Met Ser Arg  
                                  85                      90                      95  
 Leu Glu Gln Thr Pro Pro Leu Ala Val Leu Leu Ala Glu Asp Leu Gly  
                                  100                      105                      110  
 Ala Arg Ala Ala Ala Gly Leu Asp Val Ser Gly Ala Cys Ala Gly Phe  
                                  115                      120                      125  
 Cys His Ala Leu Ala Leu Ala Ser Asp Ala Val Arg Ala Gly Ser Ala  
                                  130                      135                      140  
 Arg His Val Leu Val Val Gly Thr Glu Arg Met Thr Asp Leu Val Glu  
 145                      150                      155                      160  
 Arg Ala Asp Arg Thr Val Ser Val Leu Phe Ala Asp Gly Ala Gly Ala  
                                  165                      170                      175  
 Ala Val Val Gly Pro Ser Ala Arg Pro Gly Ile Ser Pro Pro Ala Arg  
                                  180                      185                      190  
 Gly Ala Ala Gly Arg Tyr Ala Gly Ala Leu Arg Met Asp Arg Gly Trp  
                                  195                      200                      205  
 Asp Thr Phe Ala Ala Asp Pro Ser Leu Gly Arg Pro Trp Met Arg Met  
                                  210                      215                      220  
 Asp Gly Arg Arg Val Phe Arg Trp Ala Met Asp Glu Val Thr Pro Arg  
 225                      230                      235                      240  
 Ala Ala Glu Leu Leu Arg Glu Ser Gly Ile Glu Pro Glu Ala Leu Asp  
                                  245                      250                      255  
 Ala Phe Val Pro His Gln Ala Asn Leu Arg Met Ile Glu Leu Met Ala  
                                  260                      265                      270  
 Glu Arg Leu Gly Leu Pro Glu Arg Thr Ala Val Ala Arg Asp Val Val  
                                  275                      280                      285  
 Arg Ala Gly Asn Thr Ser Ala Ala Ser Val Pro Leu Ala Leu Glu Ala  
                                  290                      295                      300  
 Leu Leu Asp Ser Gly Glu Val Gly Ser Gly Asp Arg Ala Leu Leu Val

Met	Arg	Ser	Glu	Gly	Gly	Thr	Glu	Glu	Glu	Gly	Ala	Pro	Val	Val	Ala
1				5					10					15	
Leu	Leu	Leu	Pro	Gly	Gln	Gly	Ala	Gln	Arg	Ala	Arg	Met	Ala	Ala	Gly
			20					25					30		
Leu	His	Gly	Val	Val	Pro	Glu	Phe	Thr	Thr	Ala	Val	Glu	Glu	Cys	Phe
			35				40					45			
Ala	Val	Trp	Gly	Thr	Trp	Gly	Glu	Glu	Leu	Arg	Ala	Arg	Trp	Leu	Asp
			50			55					60				
Gly	Ala	Gly	Gly	Glu	Glu	Ala	Leu	Glu	Arg	Ala	Ala	Val	Ala	Gln	Pro
65					70					75				80	
Leu	Leu	Phe	Ala	Val	Gly	Tyr	Gly	Leu	Gly	Arg	Ala	Leu	Gly	Ala	Gly
				85						90				95	
Ala	Gln	Gly	Ala	Pro	His	Leu	Leu	Leu	Gly	His	Ser	Val	Gly	Glu	Leu
			100						105				110		
Ala	Ala	Ala	Ala	Leu	Ala	Gly	Val	Cys	Ala	Pro	Gly	Ala	Ala	Leu	Arg
			115					120				125			
Leu	Leu	Ala	Glu	Arg	Asp	Ala	Val	Leu	Arg	Ala	Ala	Pro	Ser	Gly	Gly
			130			135					140				
Met	Leu	Ala	Val	Ala	Ala	Pro	Val	Asp	Asp	Leu	Arg	Pro	Tyr	Val	Gly
145					150					155				160	
Ala	Asp	Val	Val	Val	Gly	Ala	Val	Asn	Gly	Pro	Arg	Gln	Thr	Val	Leu
			165						170					175	
Cys	Gly	Pro	Glu	Ala	Pro	Leu	Arg	Ala	Val	Ala	Arg	Arg	Leu	Ala	Asp

```

          180          185          190
Asp Gly Leu Thr Ala Arg Arg Leu Gln Ala Asp Val Pro Phe His Ser
          195          200          205
Pro Ala Leu Ala Gly Ala Ala Arg Arg Leu Thr Arg Ala Ser Ala Glu
          210          215          220
Arg Val Ala Arg Trp Arg Pro Pro Ala Val Pro Leu Trp Ser Gly Arg
          225          230          235          240
Thr Gly Arg Ala Leu Thr Pro Gly Glu Ala Val Arg Ala Ala Phe Trp
          245          250          255
Cys Gly Gln Leu Ala Ala Pro Val Leu Tyr Trp Pro Ala Leu Gly Asn
          260          265          270
Leu Leu Ala Thr Ala Thr Ala Thr Thr Thr Thr Gly Gly Arg
          275          280          285
Gly Val Val Leu Leu Asp Ala Ser Pro Asp Gly Ser Leu Gly Ala Pro
          290          295          300
Ala Arg Arg His Pro Ala Val Arg Ser Gly Ala Ala Arg Val Val Arg
          305          310          315          320
Leu Leu Pro Ala Arg Pro Gly Asp Pro Ala Asp Asp Val Arg Ala Phe
          325          330          335
Arg Ala Ala Leu Gln Gln Ala Gly Gln Val Val Arg Asp Gly Gly
          340          345          350

```

【0080】配列番号：12：

配列の特性：

配列の長さ：425アミノ酸

配列の型：アミノ酸

鎖の数：一本鎖

トポロジー：直鎖状

配列の種類：タンパク質

ハイボセティカル：NO

アンチセンス：NO

配列

```

          Met  Asn  Arg  Gln  Val  Ala  Val  Thr
Gly  Ile  Gly  Val  Val  Ala  Pro  Gly
          1          5
          10          15
          Gly  Ile  Gly  Arg  Lys  Pro  Tyr  Trp
Glu  Gln  Leu  Thr  Ser  Gly  Arg  Thr
          20
          25          30
          Ala  Thr  Arg  Ala  Ile  Ser  Phe  Phe
Asp  Ala  Ser  Pro  Phe  Arg  Ser  Arg
          35          40
          45
          Ile  Ala  Ala  Glu  Val  Asp  Phe  Asp
Pro  Ala  Ala  Ala  Gly  Leu  Ser  Pro
          50          55
          60
          Arg  Glu  Val  Arg  Arg  Met  Asp  Arg
Ala  Ala  Gln  Phe  Ala  Val  Val  Ser
          65          70
          75          80
          Ala  Arg  Glu  Ser  Leu  Ala  Asp  Ser
Gly  Leu  Asp  Val  Ala  Asp  Leu  Asp
          85

```

	90					95		
	Pro	His	Arg	Ile	Gly	Val	Ser	Ile
Gly	Ser	Ala	Val	Gly	Gly	Thr	Thr	
				100				
105					110			
	Ser	Leu	Glu	Arg	Glu	Tyr	Leu	Ala
Leu	Ser	Asp	Ser	Gly	Arg	Gln	Trp	
			115					120
				125				
	Glu	Leu	Asp	Leu	Ser	Tyr	Leu	Ser
Pro	His	Leu	Tyr	Asp	Ala	Phe	Thr	
			130				135	
				140				
	Pro	Ser	Ser	Leu	Ala	Arg	Glu	Val
Ala	Gly	Val	Ile	Gly	Ala	Glu	Gly	
	145					150		
		155				160		
	Pro	Ala	Ala	Val	Val	Ser	Thr	Gly
Cys	Thr	Ser	Gly	Ile	Asp	Ser	Leu	
					165			
	170				175			
	Gly	His	Ala	Arg	Asp	Leu	Ile	Ala
Glu	Gly	Ser	Ala	Asp	Val	Val	Leu	
				180				
185					190			
	Ala	Gly	Gly	Thr	Asp	Thr	Pro	Ile
Ser	Pro	Ile	Ala	Val	Ala	Cys	Phe	
			195					200
				205				
	Asp	Ala	Ile	Lys	Ala	Thr	Ser	Pro
Ser	Asn	Asp	Asp	Pro	Ala	His	Ala	
		210					215	
			220					
	Ser	Arg	Pro	Phe	Asp	Arg	Glu	Arg
Asn	Gly	Phe	Val	Leu	Ala	Glu	Gly	
	225					230		
		235				240		
	Ala	Ala	Val	Leu	Val	Leu	Glu	Glu
Leu	Gly	His	Ala	Arg	Ala	Arg	Asp	
					245			
	250				255			
	Ala	His	Val	Tyr	Ala	Leu	Val	Ser
Gly	Tyr	Ala	Thr	Arg	Cys	Asn	Ala	
				260				
265					270			
	Tyr	His	Met	Thr	Gly	Leu	Thr	Pro
His	Gly	Arg	Glu	Met	Ala	Glu	Ala	
			275					280
				285				
	Ile	Arg	His	Ala	Leu	Ala	Glu	Ser

```

Gly Thr Asp Pro Ala Ala Val Asp
      290                      295
                        300
      Tyr Val Asn Ala His Gly Ser Gly
Thr Lys Gln Asn Asp Arg His Glu
      305                      310
                        315                      320
      Thr Ala Ala Phe Lys Ala Thr Leu
Gly Glu Arg Ala Arg Ser Val Pro
                        325
      330                      335
      Val Ser Ser Ile Lys Ser Met Ile
Gly His Ser Leu Gly Ala Ile Gly
                        340
345                      350
      Ser Leu Glu Ile Ala Ala Trp Ala
Leu Ala Met Glu Tyr Gly Val Val
                        355                      360
                        365
      Pro Pro Thr Ala Asn Leu Asp Thr
Pro Asp Pro Glu Cys Asp Leu Asp
                        370                      375
                        380
      Tyr Val Pro His Glu Ala Arg Glu
Thr Arg Val Asp Arg Val Leu Ser
      385                      390
                        395                      400
      Val Gly Ser Gly Phe Gly Gly Phe
Gln Ser Ala Met Val Leu Thr Arg
                        405
      410                      415
      Asp Thr Gly Ala Arg Leu Pro Thr
Ala
                        420
425

```

【0081】配列番号:13:

配列の特性:

配列の長さ:426アミノ酸

配列の型:アミノ酸

鎖の数:一本鎖

トポロジー:直鎖状

配列の種類:タンパク質

ハイボセティカル:NO

アンチセンス:NO

配列

```

Met Thr Thr Ala Pro Ser Arg Thr Ala Gln Gly Ala Pro Pro Gly Ala
1          5          10          15
Ala Leu Pro Pro Val Phe Thr Gly Ile Gly Val Ala Ala Pro Asn Gly
20          25          30
Leu Gly Thr Glu Glu Trp Trp Ala Ala Thr Leu Arg Gly Glu His Gly
35          40          45
Leu Arg Pro Val Thr Glu Tyr Asp Ala Ser Gly His Pro Gly Gly Leu
50          55          60
Val Gly Arg Val Pro Asp Phe Asp Ala Ala Arg His Leu Pro Gly Arg

```

```

65          70          75          80
Leu Leu Pro Gln Thr Asp Arg Val Thr Arg Leu Ala Leu Val Ala Ala
          85          90          95
Asp Glu Ala Leu Lys Asp Ala Ala Val Asp Pro Ala Arg Leu Pro Glu
          100          105          110
Tyr Gly Ala Ser Ala Val Thr Ser Asn Ala Thr Gly Gly Phe Glu Phe
          115          120          125
Thr His Arg Glu Ile Arg Lys Leu Trp Thr Glu Gly Pro Ala Arg Val
          130          135          140
Ser Val Tyr Glu Ser Phe Ala Trp Phe Tyr Ala Val Asn Thr Gly Gln
145          150          155          160
Ile Ser Ile Arg His Gly Met Arg Gly Pro Gly Ala Val Val Val Ala
          165          170          175
Asp Gln Ala Gly Gly Leu Asp Ala Leu Gly Gln Ala Arg Arg Val Leu
          180          185          190
Arg Lys Gly Gly Val Leu Ala Val Ser Gly Gly Val Glu Ser Ala Leu
          195          200          205
Asp Pro Trp Gly Leu Ala Ala His Ala Ser Ser Gly Thr Leu Ser Arg
210          215          220
Ser Gly Asp Pro Ala Thr Ala Tyr Leu Pro Phe Asp Arg Arg Ala Leu
225          230          235          240
Gly Thr Val Val Gly Glu Gly Gly Ala Leu Leu Thr Leu Glu Thr Pro
          245          250          255
Arg His Ala Glu Glu Arg Asp Ala Pro Arg Ile Tyr Gly Glu Leu Ala
          260          265          270
Gly Tyr Ala Ala Thr Phe Asp Pro Pro Ala Gly Ser Gly Arg Pro Pro
          275          280          285
Gly Leu Glu Arg Ala Ala Arg Leu Ala Leu Ala Asp Ala Gly Leu Ala
290          295          300
Pro Gly Asp Val Asp Val Val Phe Ala Asp Ala Ala Gly Leu Pro Ala
305          310          315          320
Ala Asp Ala Ala Glu Ala Ala Ala Leu Arg Ala Leu Phe Gly Pro Gly
          325          330          335
Gly Val Pro Val Ser Val Pro Lys Thr Gln Thr Gly Arg Leu Ala Ser
          340          345          350
Gly Gly Pro Ala Leu Asp Val Ala Ala Ala Leu Leu Ala Leu Arg Asp
          355          360          365
Gly Leu Val Pro Pro Ala Val His Leu Asp Glu Val Asp Pro Ala Tyr
          370          375          380
Gly Leu Asp Leu Val Arg Asp Thr Pro Arg Ala Leu Pro Leu Arg Thr
385          390          395          400
Ala Leu Val Leu Ala Arg Gly His Gly Gly Phe Asn Ala Ala Val Val
          405          410          415
Val Arg Gly Arg Arg Arg Pro Arg Thr Ala
          420          425

```

【0082】配列番号：14：

配列の特性：

配列の長さ：83アミノ酸

配列の型：アミノ酸

鎖の数：一本鎖

トポロジー：直鎖状

配列の種類：タンパク質

ハイポセティカル：NO

アンチセンス：NO

## 配列

```

      Met Ser Ala Leu Thr Val Asp Asp
Leu Lys Lys Leu Leu Ala Glu Thr
      1           5
      10           15
      Ala Gly Glu Asp Asp Ser Val Asp
Leu Ala Gly Glu Leu Asp Thr Pro
      20
25           30
      Phe Val Asp Leu Gly Tyr Asp Ser
Leu Ala Leu Leu Glu Thr Ala Ala
      35           40
      45
      Val Leu Gln Gln Arg Tyr Gly Ile
Ala Leu Thr Asp Glu Thr Val Gly
      50           55
      60
      Arg Leu Gly Thr Pro Arg Glu Leu
Leu Asp Glu Val Asn Thr Thr Pro
      65           70
      75           80
      Ala Thr Ala

```

【0083】配列番号：15：

配列の特性：

配列の長さ：160アミノ酸

配列の型：アミノ酸

鎖の数：一本鎖

トポロジー：直鎖状

配列の種類：タンパク質

ハイボセティカル：NO

アンチセンス：NO

## 配列

```

Met Thr Thr Thr Thr Pro Pro Asp Asp Val Arg Ala Gly Ser Leu Pro
1           5           10           15
Gly Asp Ala Ala Arg Ser Ala Ala Leu Tyr Thr Glu Val Gln Ala Phe
      20           25           30
Tyr Ala Arg Gln Ala His His Leu Asp Ala Val Arg Ala Glu Glu Phe
      35           40           45
Ala Ala Thr Phe Ala Ala Glu Gly Val Phe Ala His Ser Pro Asp Thr
      50           55           60
Pro Ala Ala Arg Gly Arg Ala Ala Ile Ala Glu Glu Val Arg Gly Phe
65           70           75           80
Asn Ala Arg Arg Phe Ala Asp Asp Pro Val Gln Arg Arg His Trp Phe
      85           90           95
Ser Met Leu Asp Val Arg Pro Gly Glu Asp Gly Ala Val Glu Thr Glu
      100           105           110
Phe Tyr Ala Leu Val Val Val Thr Arg Pro Asp Ala Ala Leu Pro Val
      115           120           125
Ile Gly Pro Ser Cys Val Val Arg Asp Val Leu Val Arg Glu Gly Gly
      130           135           140
Glu Leu Arg Thr Leu Ser Arg Gln Val Thr Gln Asp Arg Thr Leu Leu
145           150           155           160

```

【0084】配列番号：16：

配列の特性：

配列の長さ：271アミノ酸

配列の型：アミノ酸

鎖の数：一本鎖

トポロジー：直鎖状

配列の種類：タンパク質

ハイボセティカル：NO

アンチセンス：NO

配列

```

Met Thr Thr Ala Ala Pro His Thr Arg Pro Gly Glu Ala Gly Thr Thr
1           5           10           15
Arg Gly Pro Ala Leu Val Thr Gly Ala Thr Arg Gly Ile Gly Leu Ala
          20           25           30
Val Ala Glu Ala Leu Val Ala Arg Gly Tyr Pro Val Val Val Cys Ala
          35           40           45
Arg Asp Ala Glu Ala Val Ala Arg Thr Val Lys Glu Leu Ala Ala Gly
          50           55           60
Gly Ala Arg Val Glu Gly Val Val Ala Asp Val Thr Asp Ala Ala Ser
65           70           75           80
Val His Glu Leu Val Ala Thr Thr Val Ala Arg Phe Gly Pro Val Glu
          85           90           95
Val Leu Val Asn Asn Ala Gly Arg Ser Gly Gly Gly Val Thr Ala Glu
          100          105          110
Leu Ser Glu Ser Leu Trp Asp Asp Val Ile Ala Thr Asn Leu Lys Ser
          115          120          125
Val Phe Leu Val Thr Arg Glu Val Leu Thr Thr Gly Gly Met Thr Gly
          130          135          140
Arg Gly Arg Gly Val Val Asn Ile Ala Ser Thr Gly Gly Lys Gln Gly
145          150          155          160
Val Val Phe Gly Ala Pro Tyr Ser Ala Ser Lys His Gly Val Val Gly
          165          170          175
Phe Thr Lys Ala Leu Gly Leu Glu Leu Ala Arg Ser Gly Ile Thr Val
          180          185          190
Asn Ala Val Cys Pro Gly Tyr Val Glu Thr Pro Met Ala Ala Gly Val
          195          200          205
Arg Arg His Tyr Ala Asp Leu Trp Asp Val Thr Glu Glu Asp Val Leu
          210          215          220
Ala Arg Phe Glu Ala Lys Ile Pro Leu Gly Arg Tyr Thr Arg Pro Asp
225          230          235          240
Glu Val Ala Ala Leu Val Asp Tyr Leu Val Thr Asp Ala Ala Ala Ala
          245          250          255
Val Thr Ala Gln Ala Leu Asn Val Cys Gly Gly Leu Gly Asn Tyr
          260          265          270

```

【0085】配列番号：17：

配列の特性：

配列の長さ：331アミノ酸

配列の型：アミノ酸

鎖の数：一本鎖

トポロジー：直鎖状

配列の種類：タンパク質

ハイボセティカル：NO

アンチセンス：NO

配列

```

Met Thr Thr Thr Thr Thr Gly His Gln Arg Pro Gly Ser Ala Glu His
1           5           10           15
Ser Ala Arg Leu Ala Ala Pro Pro Ala Ser Ala Tyr Glu Leu Val Ala
          20           25           30
Asp Val Thr Arg Trp Pro Leu Leu Phe Thr Pro Cys Leu His Ala Glu
          35           40           45

```



```

Val Leu Glu Ser Gly Pro Gly Thr Glu Arg Val Arg Leu Trp Ala Leu
  50                      55                      60
Thr Gly Glu Gln Val Arg Gly Trp Thr Ser Arg Arg Thr Leu Asp Ser
 65                      70                      75                      80
Glu Gly Leu Arg Val Gly Phe Arg Gln Glu Asp Ser Ala Pro Pro Leu
                      85                      90                      95
Ala Ala Met Gly Gly Glu Trp Arg Phe Thr Glu Glu Gly Glu Asp Thr
                      100                      105                      110
Arg Ala Val Leu Ala His Asp Trp Thr Leu Thr Glu Pro Gly Ala Ala
                      115                      120                      125
Pro His Arg Trp Val Thr Glu Thr Leu Asp Arg Asn Ser Thr Ala Glu
                      130                      135                      140
Ile Gly Ala Val Thr Ala Trp Ala Ala Arg Thr His Ala Ala Gly Gly
 145                      150                      155                      160
Ala Asp Ala Leu Leu Phe Ser Phe Thr Asp Ser Leu Asp Ile Ala Ala
                      165                      170                      175
Pro Ala Pro Asp Val Tyr Ala Phe Leu Asp Ala Ala Asp Gln Trp Pro
                      180                      185                      190
Ala Arg Leu Pro His Val Ser Arg Val Ala Phe Ser Thr Thr Pro Ala
                      195                      200                      205
Thr Pro Leu Thr Ala Gly Ala Glu Val Gln His Leu Glu Met Glu Thr
                      210                      215                      220
Arg Ala Asp Asp Gly Thr Arg His Leu Thr Arg Ser Ile Arg Leu Gly
 225                      230                      235                      240
Phe Ala Gly Arg Leu Leu Val Tyr Lys Gln Thr Thr Leu Pro Ala Pro
                      245                      250                      255
Leu Leu Gly His Ala Gly Ser Trp Ala Leu Glu Pro Leu Pro Gly Gly
                      260                      265                      270
Gly Thr Arg Val Thr Ala Arg His Arg Val Ala Leu Asp Pro Asp Ala
                      275                      280                      285
Val Thr Glu Arg Phe Gly Ala Gly Thr Thr Leu Ala Ala Ala Arg Asp
                      290                      295                      300
Thr Val Arg Ala Leu Leu Gly Gly Asn Ser Arg Arg Thr Leu Glu Ala
 305                      310                      315                      320
Ala Arg Ala His Thr Glu Ser Ala Gly Glu Arg
                      325                      330

```

【 0 0 8 6 】配列番号 : 1 8 :

配列の特性 :

配列の長さ : 3 1 4 アミノ酸

配列の型 : アミノ酸

鎖の数 : 一本鎖

トポロジー : 直鎖状

配列の種類 : タンパク質

ハイポセティカル : NO

アンチセンス : NO

配列

```

Met Thr Thr Thr Gly Pro Ala Thr Pro Leu Ala Pro Gly Pro Ala Ser
 1                      5                      10                      15
Ala Glu Thr Val Ala Leu Ala Asp Gly Val His Ser Trp Leu Gln Pro
                      20                      25                      30
Asp Gly Gly Trp Cys Val Ser Asn Ala Gly Ile Leu Leu Ala Pro Asp
                      35                      40                      45
Arg Val Ala Leu Val Asp Thr Ala Ala Thr Glu Ala Arg Ser Arg Ala
 50                      55                      60

```

```

Leu Gly Ala Ala Val Ala Gly Leu Ser Pro His Pro Val Arg Leu Leu
65          70          75          80
Val Asn Thr His Phe His Gly Asp His Ser Phe Gly Asn Gly Ile Leu
          85          90          95
Gly Lys Asp Ala Val Ile Val Ala His Glu Arg Thr Arg Thr Glu Met
          100         105         110
Ala Glu Ala Gly Leu Gly Leu Thr Gly Leu Trp Pro Gly Val Asp Trp
          115         120         125
Gly His Val Asp Pro Val Leu Pro Gln Leu Thr Tyr Arg Arg Arg Leu
          130         135         140
Thr Leu His His Gly Asp Leu Arg Val Glu Leu Ile His Pro Gly Pro
145         150         155         160
Ala His Thr Thr Asn Asp Thr Leu Val Trp Leu Pro Glu Gln Arg Val
          165         170         175
Leu Phe Ala Gly Asp Val Leu Leu Pro Gly Ala Thr Pro Phe Val Leu
          180         185         190
Met Gly Ser Val Thr Gly Ser Leu His Thr Leu Arg Leu Leu Arg Arg
          195         200         205
Leu Gly Pro Arg Val Val Val Gly Gly His Gly Pro Leu Ala Gly Pro
          210         215         220
Glu Val Ile Glu Glu Thr Glu Arg Tyr Leu Leu Arg Leu Arg Arg Ile
225         230         235         240
Ala Thr Glu Gly His Ala Ala Gly Leu Thr Pro Leu Glu Ala Ala Arg
          245         250         255
Arg His Gly Pro Gly Pro Phe Ala His Trp Ser Glu Pro Glu Arg Leu
          260         265         270

Ala Ala Asn Leu His Arg Ala Tyr Ala Glu Leu Gly Pro Ala Pro Leu
          275         280         285
Gly Thr Pro Leu Asp Val Leu Ala Cys Phe Gly Asp Leu Ile Ala Tyr
          290         295         300
Asn Asp Gly Glu Leu Pro Val Ser His Ala
305          310

```

【0087】配列番号：19：

配列の特性：

配列の長さ：305アミノ酸

配列の型：アミノ酸

鎖の数：一本鎖

トポロジー：直鎖状

配列の種類：タンパク質

ハイポセティカル：NO

アンチセンス：NO

配列

```

Met Thr Gly Arg Val Thr Gly Thr
Met Ala Gly Pro Leu Pro Gly Thr
1          5
10          15
Val Leu Val Ala Gly Ala Ser Gly
Ala Val Gly Gly Arg Leu Val Ser
          20
25          30
Arg Leu Leu Asp Ser Gly Val Pro
Val Arg Ala Leu Val Arg Ser Ala
          35          40

```

				45					
	Ala	Arg	Gly	Arg	Ala	Leu	Ala	Ala	
Val	Gly	Ala	Glu	Thr	Val	Val	Gly		
		50					55		
			60						
	Asp	Leu	Ala	Asp	Thr	Ala	Ser	Leu	
Ala	Ala	Ala	Leu	His	Gly	Val	Glu		
	65					70			
		75					80		
	Arg	Ala	Phe	Leu	Leu	Leu	Gln	Asp	
Glu	Ala	Gly	Ala	Pro	Phe	Ala	Arg		
					85				
	90					95			
	Ala	Ala	Ala	Lys	Ala	Arg	Gly	Leu	
Arg	Glu	Val	Val	Val	Leu	Ser	Ala		
				100					
105					110				
	Thr	Ala	Ala	Ala	His	Pro	Glu	Tyr	
Asp	Asn	Pro	Met	Phe	Gln	Lys	His		
			115					120	
			125						
	Val	Arg	Gly	Glu	Arg	His	Val	Arg	
Gly	Ser	Gly	Val	Pro	Trp	Val	Phe		
		130					135		
		140							
	Leu	Arg	Pro	Gly	Ala	Phe	Ala	Ser	
Leu	Ala	Leu	Leu	Trp	Ala	Pro	Ala		
	145					150			
		155					160		
	Leu	His	Gly	Asp	Gly	Val	Val	Arg	
Val	Pro	His	Ala	Gly	Leu	Ala	Val		
					165				
	170					175			
	Pro	Leu	Ile	Asp	Pro	Arg	Asp	Ile	
Ala	Asp	Val	Ala	Ala	Ala	Val	Leu		
				180					
185					190				
	Thr	Ala	Pro	Val	Glu	Arg	Trp	Ala	
Gly	Lys	Ala	Val	Pro	Leu	Ser	Gly		
			195					200	
			205						
	Pro	Glu	Val	Leu	Thr	Leu	Pro	Gly	
Arg	Thr	Ala	Val	Leu	Ala	Ala	Glu		
		210					215		
		220							
	Leu	Gly	Arg	Pro	Leu	Arg	Val	Glu	
Pro	Val	Pro	Glu	Glu	Glu	Trp	Val		
	225					230			
		235					240		
	Gln	Leu	Thr	Ala	Arg	His	Ile	Pro	

```

Glu Pro Tyr Ala Arg Ala Leu Ala
245
      250      255
      Gly Val Glu Arg Phe Phe Thr Glu
His Pro Pro Ala Val Ser Pro Gly
260
265      270
      Val Thr Asp Val Thr Gly Arg Pro
Ala Arg Ser Phe Ala Thr Trp Val
275      280
      285
      Arg Asp His Ala Ala Ser Phe Ala
Pro Gly Pro Pro Thr Val Val Gly
290      295
      300
      Gly
305

```

【0088】配列番号：20：

配列の特性：

配列の長さ：396アミノ酸

配列の型：アミノ酸

鎖の数：一本鎖

トポロジー：直鎖状

配列の種類：タンパク質

ハイポセティカル：NO

アンチセンス：NO

## 配列

```

Met Thr Glu Gln Ser Thr Thr Leu Ala Ala Pro Ala His Ala Glu Arg
1      5      10      15
Thr Gly Arg Arg Thr Lys Gly Pro Gly Arg Pro Pro Asp Arg Ser Arg
20      25      30
Asn Gly Pro Gly Arg Ser Pro Arg Ser Arg Pro Gly Thr Arg Ala Ala
35      40      45
Val Arg Arg Ser Gly Val Ser Pro Gly Lys Ser Ser Ala Gly Ile Val
50      55      60
Ala Ala Gly Phe Ala Arg His Phe Val Pro Ala Ala His Gly Gly Ala
65      70      75      80
Glu Gly Gly Phe Gly Glu Leu Ala Glu Ala Val Leu Arg Leu Gly Thr
85      90      95
Gly Cys Thr Ser Ala Ala Trp Ala Ala Ser Leu Ser Ala Tyr Ala Gly
100      105      110
Arg Tyr Ala Ala Tyr Leu Pro Glu Glu Gly Gln Ala Glu Val Trp Ala
115      120      125
Glu Gly Pro Asp Ala Leu Leu Ala Gly Ala Leu Val Pro Ser Gly Thr
130      135      140
Val Thr Pro Val Pro Gly Trp Arg Leu Asp Gly Ala Trp Pro Tyr
145      150      155      160
Ile Ser Gly Val Arg His Ala Ala Trp Val Leu Ala Cys Ala Thr Val
165      170      175
Pro Gly Gly Glu Gly Glu Glu Gly Pro Glu Val Arg Phe Phe Ala Gly
180      185      190
Pro Arg Ala Ala Pro Arg Val Glu Arg Thr Trp Asn Thr Thr Gly Met
195      200      205
Arg Ala Thr Gly Ser Asp Thr Leu Val Leu Asp Asp Val Leu Val Pro

```

```

210          215          220
Ala His Arg Ser Phe Pro Arg Thr Arg Val Leu Ala Gly Gln Arg Pro
225          230          235          240
Arg Val Ala Gly Ala Val Pro Thr Val Arg Met Ala Arg Val Gly Ala
          245          250          255
Leu Pro Val Val Thr Pro Leu Val Gly Ala Ala Arg Gly Ala Leu Arg
          260          265          270
Ala Trp Thr Glu Arg Ala Ala Gln Gly Arg Ala Pro Ser Pro Gly Ala
          275          280          285
Leu Gly Glu Leu Ser Arg Ala Ala Gly Glu Ala Asp Ala Ala Glu Leu
          290          295          300
Leu Val Leu Arg Ala Ala Ala Ala Ala Asp Gly Thr Val Ser Leu Pro
305          310          315          320
Glu Pro Ala Ala Ala Val Arg Gly Lys Arg Asp Thr Ala Leu Ala Ala
          325          330          335
Glu Leu Ala Val Gly Ala Val Gln Arg Leu Val Arg Ala Ser Gly Thr
          340          345          350
Ser Gly Gln Ser Pro His Asp Pro Val Gln Arg Phe Trp Arg Asp Val
          355          360          365
Gln Thr Gly Ala Ser His Val Ala Leu Ser Pro Glu Ala Ala Gly Ala
          370          375          380
Ala Tyr Gly Ala Trp Ala Val Gln Glu Ala Asn Arg
          385          390          395

```

【0089】配列番号：21：

配列の特性：

配列の長さ：108アミノ酸

配列の型：アミノ酸

鎖の数：一本鎖

トポロジー：直鎖状

配列の種類：タンパク質

ハイポセティカル：NO

アンチセンス：NO

配列

```

Met Ser Gly Gly His Met Val Thr M
et Ile Asn Gln Met Leu Leu
1 5
10 15
His Gly Asp Glu Ser Arg Phe Leu
Ala Val Leu Glu Glu Ile Cys
20
25 30
Ala His Met Arg Ala Gln Pro Gly
Phe Leu Ser Leu Arg Leu His
35
40 45
Arg Ser Pro Asp His Pro Glu Arg
Trp Ala Met Leu Ala Asp Trp
50
55 60
Ser Asp Ala Ala Ala His Arg Ala
Ala Ala Ser Ala Pro Gly Ile
65
70 75
Arg Pro Ala Phe Ala Arg Leu Arg

```

Ala Glu Ala His Thr Ala Pro  
80  
85 90  
Gln Val Tyr Ala Pro Val Pro Thr  
Pro Gly Ala Pro Ala Gly Asp  
95  
100 105  
Ser Leu Ala

【0090】配列番号：22：

配列の特性：

配列の長さ：24379塩基対

配列の型：核酸

鎖の数：一本鎖

トポロジー：直鎖状

配列の種類：cDNA

ハイボセティカル：NO

アンチセンス：NO

## 配列

CAATTGCCAT GGAATGCCCG GTTATAAGGG AAGTGGAAATG TGTCACCGCC GTCGTCGCGT 60  
GGGGCGGTGG GCGGCTAATG TGAGCGATCC CGGTGGCCGC GACCGTACTC ATCGCGGCGC 120  
GTCACCGATT GACCCGAGTG CCTCCCGCCG CGTCTCCACC CCACCTGGAG GCGGGGTCGG 180  
CGCCGCGTGC CCCGTGACGG AGGATTTGGA CAACCGTGA CGCGCAAGGC CGTGGGCGGG 240  
GGGAAGAATT CGACCCCGCA GACCAATGGG TACTCAATCC GCAGACCGGC GAATACGAAC 300  
TGCGACTCGG CGGCTCCGCT GAGCAGTCGC AGGAGCCGCC GGGGTCCGGC GCGGGGGCCT 360  
GGGACCACGT CGACGGGTCC GCGCGTCCCC GCAGAACCGA GCGTACCGC TCCGAGCCCC 420  
GGGTGTCCGA GGCAGCGGCG GCTCCCGCG CACGAGAAG GCGGCGCGC CATGGCGGTC 480  
GCGCGGCGGC GCAGTAGGAG GCGGCGGCCG GCGGTCTGCC CTCGGCCCGC AAACAGGCTT 540  
CCGGGCGGGC CGCCTCGCGC AAGGCGCGCA AGGCGGCGG CAAGAAGAAG GCGCTGCTGT 600  
GGACGGCCGG GGGGCTCGGT TTCGCGCTCG TCGCGGTGGC CGGCAGCGGC TACGCGTACT 660  
ACCAGCACCT CAGCGGCAAC ATCGACAAGA TTGACGTCCG CGACGCGGC AACAAGGACG 720  
CCGCCCCCGA CGGGCCGATC AACATCCTGA TCATCGGCAC CGACAAGCGC ACCGGCAAGG 780  
GCAACGAGGG GTGCGGGGGC AAGGACAGCC CTGGCCACGC GGACACCAAC ATCTTGCTGC 840  
GCGTCTCCGC CGACCGCACC AACACGACCG GCCTGAGCAT TCGCGTGAC CTCATCACGA 900  
ACATCCCGGA CTGCTGACC ACCCAGGACG ACGGTTGAA GAAGACGATC CCCGGCACGC 960  
AGAACGTCCG CTTCAACACG AGCCTCGGGC AGGAGGGCCG CGACCCCGC TGCACGATGC 1020  
GCACGGTCAC CGAGCTGACC GGGCTCAAGG TCGACCACTT CATGATGGCC GACTTCAACG 1080  
CGGTCAAGAA CCTGACGACG GCCGTGAACG GCGTCGAGGT GTGCGTGGCG AAGGACGTGG 1140  
ACGACCCCGA CTCGCACCTC AAGCTCTCGG CGGGACGCA CAAGGTGCAG GCGGAGCAGG 1200  
CGCTCGCCTT CGTCCGCACC CGGCACAGCT TCGGCAACCA GGGCGACCTC GACCGCATCA 1260  
AGGTCCAGCA GCAGTTCCTC GGTTCGCTCG CGCGGACGCT CAAGTCCGAG GACACGCTGA 1320  
CGAGTCCGAA GAAGCTCTAC AAGGTCGCG AGGCGGCGAC CGACGCGCTG ACCGTGGACT 1380  
CGGGGATCGG CAGCATCACG AAAGTATGATG CGCTGGCCAA GGAGTTGCAG CACATCAATC 1440  
CGAAGAACAT CACCTTCGTG ACGCTCCCCG TGGTCGACAA CCCGGCCGAG AAGGTCAAGG 1500  
CGACGGTCGT CCTCAATGAG ACGGACGCTG ATCGCGAGCA GTCGGCTCTC GGGCAGAGTC 1560  
TCGACGTGGG CCGCCAGCTC GTCGACTCGC TCACCGACCA GGACCCGCGC GACGGCAAGA 1620  
CCGTCCCGTG GCTCGGACG CGCTGGAAGG CCGACCCCGA GGCCACCGC TTCACCTTCA 1680  
CCCTGCGGC GGGGCCACT TTCAGCGACG GCACACCGGT CGACGCGCGC GCGGTCAAGG 1740  
CCAACCTTGA CGCCGTGCAC GAGCTCGGCG CCGCCGCCTC GCGGGGGGCC GTGTACCTCG 1800  
ACGGCTACCG CGAGACGCGC GTCAGCGGCG CCCGTACGCT CACGGTCGTC TTCGACAAGC 1860  
CGAACGCCCC GTTCTTGCGC GGCACGTCCA CCGTCTCACT GGGCCTGCTC TCGCCGGGCA 1920  
GCCTGCGCG CACCCGCGAG GAGCGCTGCA CCGGGCGCCT CGTCGGCTCC GGTCCCTTCG 1980  
TGCTCGACCG CTACCGGCG AACACCTCCG TCACGCTGGA CCGGAGAAAG GGCTACTCCT 2040  
GGGGCTCGC CCGTGGGCAA CCGGAGGGCG GCGCGTACCT GGAGGGGGTC GAGTACCGGA 2100  
TCGTGCCCCA GAACACGACC CGCTCGGGCG CCCTGTCCGC CCGTCAGCTC GACGTGCGCA 2160

CCGCGCTCGC	GCCGCAGGAC	AGGGAGCGGT	TCTCCGCCCC	CGGCTGGTCG	CTGCTGACCC	2220
GTACCGCGCC	CGGCGTCGAC	CTCAGCCTGT	ACGTCAACGC	GCGCCGCACG	GCGCTGCGCG	2280
AGGCGGCGGT	GCGCCAGGCC	CTCCAGAAGG	GCATCGACCG	CGAAGCCGTC	GCCACGACCT	2340
TCCTCAGTTC	CCGCAAGCTC	GCGGCGACCA	GCGTGCTCTC	CTCCACGACC	CCCGGCTACA	2400
CCGACCTCGG	CGACCGCCTC	GCCCACGACC	CTGCGGGTGC	GAGGCGGCTG	CTCGACAAGG	2460
CCGGATGGCG	GCCCGGGGCG	GACGGCATCA	GGGTCAAGAA	CGGCGTCAGG	CTGCGACTCG	2520
ACGCCGTCTT	CGTCCGGCAG	CAGAGCCTCG	AACTCGTGCA	GCAGCAGCTC	AAGGACATCG	2580
GCGTCGAACT	GCGGCTGAGA	CAGCTCACCG	TCTCCCGTTT	CCCCGAAGTG	CTCGCCGCGG	2640
GCAGCTACGA	CCTGAGCCTC	CAGAGCGCCA	ACCGCGCCGA	TCCCGACGTC	CTCACCACGG	2700
CCTTCGCGGG	CGGCACCCCC	GTGCGCGACG	CGCGCCTGCG	CTCCGAAGTG	CGCCGGGCCA	2760
CCTCTCCAC	CGACGAGGCG	ACGCGCTCCT	CGCTCTTCGC	CGCGGCGCAA	CGGCGTCTCG	2820
TCGACGAGGG	GCACGTCCTG	CCGCTCAACG	AGACCGAGGA	GACCGCCGCG	CTCTCGACGC	2880
GGGTGCACGG	GCTGACGCGG	GACGCCTCCA	ACCGGCTGGT	CCTGCACGAC	ACCTGGACCA	2940
CCGGGTGACG	GCCCGCTACC	TGGCCCGCCG	GCTGGGCCCG	GTGCTCCTCG	TGCTCTGGGC	3000
CGCGTACACC	CTGTCCTTCG	CCGTCTCTA	CCTGCTGCCG	GGAGATCCCG	TGCAGACCAT	3060
GCTGAGCGGC	GCGGCGGGCG	GGGACGGGGC	GGCCGTGAC	CCGCAAGAGG	CCCAGCGGCT	3120
GCGCCACACG	CTCGGCCTCG	ACCGCCCGCT	CGCGTCCAG	TACACCAGCA	TGCTGGGGCA	3180
CGCCCTGGCG	GGAGACCTCG	GCACCTCGAT	CCGACGCGGG	GCGCCCGTAC	GCGGGCAACT	3240
CGCCCAAGCG	CTGCCCGACA	CCCTCTCGGT	GGCCCTCCCC	GCGCTCGTCC	TCAGCGTGCT	3300
CGTCCGCGTC	TGCCCTCGCC	TCCCTCGGGC	CTGGCCCCCG	CGACGAGCCC	TGCGCAGGGC	3360
GGCGACCGCC	CTGCCCTCGC	TCGGCACCGC	GATGCCACGC	TTCTGGCTGG	GGCTGCTCCT	3420
CGCCCAAGTG	GTCTCCTTCC	GCTGGGACT	GCTCCCGGCC	ACGGGCGGCG	GGCGTCCGCC	3480
CGGGGCCACC	CTGCTGGCCG	CGCTGACCCT	CGCCCTGCC	ATCGGCTGCG	TGCTCGCCCA	3540
GGTGCTGGGG	AGGGGCCTGC	GCGCGCGCT	CGCCGAGCCG	TACGCGGACG	TGGCGCGCTC	3600
GCGCGGCGCG	GGGCGGGCCA	GGCTGCTGCT	CGCGCGCGCG	CTGCGCAACG	CCTCGGTGGC	3660
CGCCCTGGCC	CTGCTCGGCG	TCGTGTGCGG	GCAACTCCTC	GCCGGGGCCG	TCCTGGTGGA	3720
GACCGTCTTC	GCCAGGGGCG	GCATCGGGCG	CCTCGCCATG	GACGCGTCA	CCTACCAGGA	3780
CCTGCCCGTC	GTGCAGGGCG	TGGTGGTGCT	CGCCGCCCTC	GTGCGGCGC	TGGTCAACCT	3840
CGTGGTGGAC	CTCCTCCTCC	CGCTCCTCGA	GCCGCGTACC	GCCTCGGAGG	CCGCCGATGC	3900
CGTCCCTGCC	CACTGACTCC	CCGGGCCCGG	CCCGTACCC	GGCCCTCCC	GCCCAGCCGG	3960
TCACGGACCG	CGCCAGGGCG	CCGGAGCCCC	GGGAGGCGGG	AGCGCGCCGC	GCGGGCCGCG	4020
TGGCGCTCCG	GCGCGTCGCC	GCGCTGTGGC	GGGCGCCCGG	ACTGGCCCTG	TCCCTGCTCG	4080
TCCTCCTGCT	CGTCTCGGCG	TGGGCGCTGC	TGCCCGGCC	GTTCACGGCG	GCCGACCCTC	4140
TGCGGGCCGA	CGCCGCGCAC	CGCCTGCTCG	CCCCCGGCGC	CGACACCCG	CTCGGCGCCG	4200
ATCACGTCCG	CCGCGACCTC	TACGCGCGGG	TCGTGCACGG	CACCGCGCGC	TCGTGGGCA	4260
CGGCGTTCCG	CGCCGTGCGG	CTCGGGGTGC	TCGCCGGCGG	TGCGCTCGGG	GCCGTGGCCG	4320
GAGTGGCGGG	GCGCGCGGTG	GACGCCGTGC	TCATGCGCGT	CGTGGACGTC	CTCCTGGCCG	4380
TCCCGGGGCT	GCTGCTCTCG	CTCGCGGTGG	TCTCCGCGCT	CGGCTTCGGC	ACCGCGCAGG	4440
TCGCTGCGC	GGTGGCGGTG	GGCACCGTCG	GCGGGATCGC	CAGGGTGAGC	CGTGGCGAGG	4500
TGCGGCGCGT	GCGCGGCGGC	GAGTACGTGC	AGGCCGCGCG	CCTCGCCCGG	GTCGCGGGGC	4560
CGCTGATCCT	GCTGCGCCAC	ATCGTGCCCA	ACGCGGCTCC	GCCCGTGCTC	GCGCTCGCCG	4620
TCACCGAGTG	CGGCACCGCG	GTAATCGGCG	TCGCGTCCCT	CGGCTTCCTC	GGCTTCGGCG	4680
CGCCGCGGCC	CGCCCGGAG	TGGGCGGCC	TCATCTCCAC	CGGCCGGGAC	TACCTCGTCT	4740
CGCCTGGTG	GCTCACCACC	CTGCGGGGCC	TCGTGCTCGT	CGCCCTCGTC	GTGCGCCTGC	4800
ACCGCGTCGG	CCGCGCCCTG	GAACGAGAGG	AACGCACCGG	ATGACCCCGG	CGGACAAGCC	4860
GACCGACGAG	CGCTCACCTG	TCCTGGACCT	CAGTGGTGTC	AGCGTCGCCT	ACGGGACGCG	4920
CACCGTGCTG	CACGGCATCG	ATCTGCGCCT	CGCGCCGGGG	CAGGTGCTCG	CGGTGCTCGG	4980
CGCCTCGGGT	TCCGGCAAGA	GCACCTCGC	GCAGGCGGCT	CTCGGCCTCC	TGCCCCCGGG	5040
CGGCGGGGTG	ACGGCAGGGC	GCGTCACGGT	CGCCGGGCAC	GACATCACCG	CGCTCGCGCC	5100
GCACCGGCTG	CGCGCGCTGC	GCGGCACGGT	CACCGGACTC	GTGCCGACAG	ACCAGGCGGT	5160

CTCCCTCGAC CCCCTCGTCC GCGTCGGGCG GCAGGTCACC GAGACGCTGC GCGCCCATCG 5220  
 CCTGGAGGAC CGCCGCGAGG CCGCCCGGCG CGCCGTGCCG TTGCTCGGCG AGGCGGGGAT 5280  
 CGAGGCCCCC GGGCCGCTCG CCGCGCCTA CCGCAGCGG CTCTCCGGCG GCCAACGCCA 5340  
 ACGCGTCCTC GTCGCGGGG CCTTCGCGGC CGCCCCCG CTGGTCGTGC CCGACGAGCC 5400  
 GACGAGCGCG CTCGACGCGA CCGTCGGGCG CCGGTCATG GACCGCTTCG CCGCCTCGT 5460  
 CGCCGCGCAC GGCACGGCGG TCCTCCTCGT CACCCACGAC TTCCGGCTCG CGAGGGAACG 5520  
 CGCCGACCAG GTCGCGCTCC TCGCGGACGG CCGTCTCGTG GAGAGCGGCC CCGCGCACG 5580  
 GGTCTCTCGAC CGGCCCGCCC ACCCGTACAC CCGCCGCCTC ACGGGGGCGG GCCGCCGCGT 5640  
 GGCCGCACGG GGCACGGCGC CGCGCGGTC GGGCACCCCC GTGGTGCGCG CCCGCGACCT 5700  
 CGTCAAGGAG TACCGGCGGG ACGGCGGGG GGTGCGCGCC GTGGACGGGG TCGGCTTCAC 5760  
 GGTCGCGGAA GGCAGGTTCT TCGCCCTCGT CGGCGAGTGG GGCTCGGGCA AGTCCACGAC 5820  
 CGCCCGGCTC GTGACGGGTC TGACCGCGCC CACGTGGGCG GCGGTGGAGC ACGCGCCGCG 5880  
 TCCGGTGCGG CCCAGCTCG TCCAGCAGAG CCCGTACGCG GCCTTCGACC CGCGCTGGAC 5940  
 CGTGCGCCGC ATCGTGGAGG AGCCGCTGCG GCGCGGGCAC GTGCCCGGGG CGCGGCGGCG 6000  
 GCGCGGGCTG CGGAACTGC TCGCGCTCGT CGGCCTCGAC GAGGAGCTGC TCGCGCGGCG 6060  
 GCCCCGGGAG CTGTCCGGCG GGCAGCGGCA GCGGTGGCG CTGCCCCGCG CGCTCGCCCC 6120  
 GGAACCGCGC CTGCTCGTGT GCGGCGAACC CGTCCCGGCC CTGACCCCG TCGCGCGCGA 6180  
 GCGCGTCGTC CACCTCCTGG AGCGGCTGCG GCGGAGCTG GCCTCACGT GCCTCTTCGT 6240  
 CTCCACGAG CTGGACGTGG TGCGCGGTT GTGCGGTGCG GTGCGCGTCA TGCGGGGCGG 6300  
 GCGGCTACTG GAGAGCGGAC CGGTGGGGGA GGTGCTGTGG GCGCCCTCGC AGCCGTACAC 6360  
 GCGGGCGCTG CTGCGCGCGG AGGCGGGCCC TTCGACACG CCGGAGCGG GGTGAGCTGC 6420  
 GTCACGTGCG GAGGGGCTTG CGCGGCCCC GCCAGCCCGC CTACGCGAAG GGAGGGGCT 6480  
 TACCTGGAGG TATGGGCTTA CCGGGGAAA GTCTGCAGCC TAGGCTCGCC CGTGAACGA 6540  
 GATCACTGTC GAGATCTGGA CCGACGTCT GTGCCGTGG TGCTACATCG GCAAGCGGCG 6600  
 CTTGAGCGG GCGCTGGCGG CCTTCGAGCG GAAGGAGGAC GTGCGGTCC ACTGGCGCAG 6660  
 CTTGAACTC GACCCGCGCG CCCTGCGGGT CACCGACGAG ACCATCCCCG AGCGCATGCT 6720  
 GCGCGGCCAG GGCATCCGCG CCGAGCAGG CCGGAACTC CTGCGCGGG TGAGCGCCCA 6780  
 GGCCGAGGCC GAAGGCTGG AGTACCACCT CGACCGCGCC CGGCCCTGCA ACACCTTCGA 6840  
 CGCGCACCAG CTCGCCACC ACGCGGGCAC GCGGGGCTC GCGGAGACCT TCCAGGAACG 6900  
 GCTGATGTGC GCCTACACG CCGAGGGCGT CTCGGTGGG GACCACCCCA CCTGCTGGC 6960  
 CCTCGCGAG GAGGCGGGG TCGACGCGG CCGCGCCGCC GAGGTGCTGG CCGGTGACGC 7020  
 GCACGCGAG GACGTACGG CCGACGAGGA CCGGCGCGC CGGCTCGGCG TCGAGGGGT 7080  
 GCGCGCCTC GTGATCGGG GACGCTGGTC CGTCTCGGG GCCCAGCCCG CGGAACTCCT 7140  
 CACCGGTCTC CTGGAACGGG CCGGACCAG AGCCGCGCG TGACGGTCAC CCGGACCGTC 7200  
 CGCGGGGCGG ACGCGGCCG GTACCGCGCC GCAGTCTCG CACCCGCTT CACCCCTCCA 7260  
 CGCCCGCACC GGTCCGTCG GCGCCGGGC CGTCCCGCC GGGGGCCGT GCCCCTTC 7320  
 GCGGTCCCAA GGAAAGCGTC GCCATGTCT CCTCCCGCC CGCCCCGCC ACCCGGGG 7380  
 TTGCCCCGA CTCGCGCCC GCTCGCGCC TCGGCTCGT GCTCCTCGTG TGCTGCCTGG 7440  
 CGCAGTTCCT CGTGACGTC AGCGTCGCA TCGTCAACGT GCGGCTCCC GACATCCAGC 7500  
 GCGGGCTCGG CTTAGCGCC GAGAGCCTCC AGTGGTCTG CAACGCCTAC ACCGTACCT 7560  
 TCGCGGCTT CTTGCTCTC GCGGACGGA TCGCGACCT CTTGCGGCG CGCCGATCT 7620  
 TCCTCGCGG TGTGCGGTC TTGCCCCTCG CGAGCCTCG CGGGGCTTA AGCCAGAACG 7680  
 CGGGGACCT CGTCCCGCG CGCGCGTGC AGGCGCTCG CGCGCGCTC ATCGCGCCA 7740  
 CCACCTCGC CGTGTCTGG ACGAGTTCA AGGACCGCA CCAGCGCCAC CGCGCTTCG 7800  
 GCGCTGGGG CGCGTCTCC GGTGCGGGG GCGCTTCGG CGCCCTCGCG GCGGCGCGC 7860  
 TCACGAGCG GTTCTCCTGG CGCTGGGTG TCTTCGTCAA CCTCCGATC GCGGTGCTCC 7920  
 TGCTCGCGG GATCGCTGG GGCATCAGG AGTTGCGGA CGCGGGCGAG GACCGCGCA 7980  
 TCGAGTCGC GGGGCCCTC ACCGTGACG TCGGTCTCT CGCGCTCGT CTCGGCATCG 8040  
 TGCAGAGCG CCCGACGGC TGGGGCTCG CCGCCACCCT CGTCCGCTC CTCGGCGGC 8100  
 TCGCCCTGCT CGGTGCCTC GTTCTCGTG AGGGGCGGT CGCGCGCAG CCGCTCATAC 8160



CGCTCGGCAT CTTCCGCTCC CGCTCCGTCG TCGCGGCCAA CGTCGTGGCG ATGACCAGCG 8220  
 GCGCCGCGCT CTTACGATG TTCTACTTCC TCACCTCTT CCTCAACCAG GTACGGGACT 8280  
 ACAGCCCGCT GCGACCGGC TTGCCTATC TCGCGCTCGC GCTCGCCATC ATGGTCGCG 8340  
 CGCAGTTCTC CGCGCGCTC GTGCGCGTCC TCGGCCCCCG TACGACCCTG CTCGTGTCCA 8400  
 TGGCTCTCAC CGCGCCGGA CTGCTCTGGC TCTCGCGGCT CACCGAGGAC TCCGGGTTCC 8460  
 CGGGCGGACT CCTCGGCCCG ACCCTGGTCG TCGGCATCGG ACAGGGCATC TCGATGTCCG 8520  
 CCTCCGCGAT CGCGGGGTC GCCGGGGTGC GCCCGCAGCA GGCCGGGCTC GCCTCCGGGC 8580  
 TGCTCAACGC GACCCGGCAG CTGCGCGGCG CCCTCGGCCT CGCGGTCTGT GCCCGCGTCG 8640  
 CCACCTCGCG CGCGACGGA CTCTCGACG GGGTCGCGCG GCCCAGCGG GAACTCGCCC 8700  
 GGCAAGCCCA GCGCTCCGG CACCGCTCT CCATCGCGT GGCCGCGCG CTGTCCGCG 8760  
 TCGGCTTGCT CGCTCGTC GCCGCCCGG GCGGCTCTCC GGCCCCACC GGCACCGCA 8820  
 CGGGCGGCGA CTCCGCGCA CCGGCTCCG CGGCGGCCCG CGCCGCCACC GGGACCACCG 8880  
 GCCCGGGCGA GATCTGAGCG CTCCCCCGG TGACCGCCCG CTCGCGCGG GCGGTCACCG 8940  
 GGGACGGCGC GGCTCGGTA TAAACATGCG CGGGCCTTT GCATGTGTGG GAAAGATGCG 9000  
 GGGCGTGAC ATAATCACCG GCTTGCGCGC CGCGACTGTG TGCTGCGTCA CGTTTCACGT 9060  
 GGTGTCTATA TTTCCCTGTC TATTCTTGGC ACCGCTCGGA GGGGACCGG TCGCGGCTGC 9120  
 CGTTTATCCC GCCGACGTC ATGATGTCCA CCGGAGGGA AACGTGAAA TCAAGTACAT 9180  
 GGGTCAGTTG ACCATGCGGT GGGAGGGCG GAAAAGCTC CCCTCGGCG GCAAGCGCG 9240  
 CACCGTGCTC GCCCTCTGC TCCTCAACGA CAAGACACC GTACGACCA GCGCACTCAT 9300  
 CACGGAATC TGGGGCGAGA ACCCGCCCG CAGCGCCCTG ACCACCCTCC AGACCTATAT 9360  
 TCTCCAGTC CGGAAATGCC TCGCGCGAT GAGCGGGCGA AGCCTCGCCT GCATATCCGA 9420  
 GAAAACCTG GTGACGTGGC CTGCGGTTA TCTGGCGCG CTTCCGGCGG ACGCCACTTC 9480  
 GGATGTGGCT GAATTCGTC GGTTCGCCAG GGAGGGCGG GAAGCGAAC GCAGGGTCA 9540  
 TCTTGCGAG GCCGTCCGT CTACCGCGC GCGGCTGTCA CTGAGCCAGG GACCACTCCT 9600  
 CGCCGATATC GAGCACGGC CGTGCTGCG CGCGAGGCG GTACGCATGG AGGAGTGCCG 9660  
 GCTTTCGTC GTGGAACGCT CCATCGAGG CGACCTGCTC CTGCGCGGC ACCGGGAGGT 9720  
 CGTCAGCGAA CTGTCCGCC TCGTCGCCA GTACCCCTAC CACGAACAG TCACGGGGCA 9780  
 GCTCATGGTC GCGCTCGTC GCTGTGGCG CCGGCAGGAC GCGCTCGCG TGCACCAACG 9840  
 GCTCCGGCC CGCATGGTG AGGACCTCG CCTCGAACG AGCAGCCACC TCGCGCCCT 9900  
 GCAGTCGCC GTCTCAGCG GCGAGCCGT GCCGGACCG CCGGAACCG GTGGGAGAT 9960  
 CCCCACCCC TACGCGGGC CTTTCGCCAC CGCGCGCGC TCGCAGACT GACCCCGCCC 10020  
 GGGGCCCGT CCACCGTCC CGTCCCGG GTGGAAGGC TGGACGAGA ACAGGCCCTC 10080  
 TCCTTCCCG CGCGGGGCC CGCGCTCAG GCAGCGCGG AAACCCCGC TCCAGTGCA 10140  
 CGAGCGGTT GCCCTACCC TGCCCCACC CGCGACCTC GCCGACGAG ATCAGTGGT 10200  
 CCCCAGCAG GTGCCGGCG TGTACCGCG ACTCCAGCT GACGAGCGG TCCGCCGCA 10260  
 CCAGCAGCCC CGAGCCCTC AGCGCCACGA GCCCCCGC CCGAAGTTG TCCGCCCGC 10320  
 GCGTCGCGA CCGGAGGCC AGCTCCCGT GTCCGCTCG CAGCAGCTC ACCGGAACC 10380  
 GCGCGCACT CTGGAACGCC CCGTAGGACT GCGCGCCGT GCCCTGGAC ACCGACCA 10440  
 GCGGGGAGC GAGCGACCC GAGCAGAAG AACTCGCGT GAAACCGCG GGGGTGCCGT 10500  
 CCTCTCGCG CGTCGTCAG AGGACCACC CGCGGGGAA GCACGACAT CCGGCGCGGA 10560  
 AGCCTCGGC ATCGACCGC CCGACCTGCC CCGTCCGCG GCGCGTCCGT GTCTCGTTCA 10620  
 CGGGAGCTC ACCACCTGG CCGGTAGTT GAGCCCGCG CCGAACCCA CGAGGAGCG 10680  
 GCGTCCCCG CTGCCACCT CGCCGAGTC GAGGAGCGC TCCAGCGCA GCGGACCGA 10740  
 GGCCGCGAG GTGTTCCCG CGCGACCAAC GTCCCGCGC ACGGCCGTG GCTCCGGCAG 10800  
 GCGGAGCGT TCGGCCATCA GCTCGATCAT GCGCAGGTT GCCTGGTGCG GGACGAAGG 10860  
 GTCCAGCGC TCCGGCTCG TCCGCTCTC CGCAGCAGT TCCGCGCCC GCGGGTCAC 10920  
 CTCGTCCATC GCCCAGCGGA ACACCGCCG TCCGTCCAT CGCATCCAG GCCTGCCGAG 10980  
 CGAGGGATC GCGCGAACG TGTCCAGCC CCTGTCCAT CGCAGCGCT CCGGCTACCG 11040  
 GCCCGCGCA CCACGCGCG GCGGCGAGT CCCCGGCCG GCGAAGGCC CCACACAGC 11100  
 GGCTCCGCC CCGTCGCGA AGAGCACGA GACGGTGGG TCCGACGCT CCACGAGATC 11160

CGTCATCCGC TCCGTGCCCA CCACCAGGAC GTGCCGGGCG CTCCCGGGCG GCACCGCGTC 11220  
 CGAAGCGAGC GCCAGCGCGT GGCAGAAACC GGCAGAGGCG CCGGACACGT CCAGCCCCGC 11280  
 CGCCGCCCGC GCGCCGAGGT CCTCGGCGAG GAGCACGGCG AGCGGGGGCG TCTGCTCCAG 11340  
 CCGCGACATG CTGCCACCA GCACCAGATC GATCCCGTCC GGCAGACCCC CGCCCGCGCG 11400  
 GAGCGCCTTC TCGGCCGCGG CCGCCCCAT GAACCGAAG GTCTCGTCCG GCGCGGCGAA 11460  
 GCCGCGCGCA CGGATACCGC TGCGGGTCTC TATCCACTCC GGGGTGACG CGACCGTGGC 11520  
 ACACAGTTGG GCGTTGTCCA CACTGCGCGG GGGCCGGTAA CCCCCACCC CCAAGAGCCT 11580  
 GCTGTGCACG ACAGGCGCGG TACCGAGCAT CTCCGCCTCG CCGGTCACTG TCCCTCCCTC 11640  
 CCGCCAGGGG CCGGTGCGCT CCGGCGGCTG ACAGCCTCGG GGGCGCTCT TGAGCCCGGA 11700  
 TCGAGCGAG ATCGAGACCC CGGCCGGATG CTGGGGCCGT CACTCCCCGG CGCCGGCGCG 11760  
 GGGCCGAGGG GACCGGGCTC CGGTCCCCA GGAAGGAACA CACGCGCAT GAGCGCCTTC 11820  
 ACTCTACCG AGTTCAAGAA GCTGGTCGAG CAGAGCTACG ACGCCGAATC GGCCGAAGCC 11880  
 CTCACGGGC AGGCCCTCGA CACGAGCTTC ACCGATCTGG GTTACGACTC GCTGACGGTC 11940  
 TACGAGATCG TCACCGCAT CCAGGACGAG CACGGAGTCA CCGTGCCCGA CGAGGAGCTG 12000  
 GACCTCCTCG ACACCCCGCG CGCCCTCATC GCGTACGTGG ACGCGCGCGC GGGTTCCTCG 12060  
 ACCTGACCCG ACGGGAGGGT GCGGGGAGGG CCGGGCCGGA GCGGGCGGGG CGCGGCGGG 12120  
 GTACGACGCC GCCGCGCCC GGGCGGCAGG CGCTCAGCCC CCGTCCCGTA CGACCTGACC 12180  
 CGCCTGTTGG AGCGCGGCGC GAAAGGCCCG CACGTCTGCC GCGGGGTGCG CGGGACGCGC 12240  
 GGGCAGGAGC CGAACGACAC GGGCCGCGCC CGAGCGGACG GCGGATGAC GCGCGCGCCG 12300  
 GGGCCCGAGA CTGCCGTCCG GCGAGGCGTC GAGCAGCACC ACGCCACGCC CGCCGCCCGT 12360  
 GGTCTGCGCC GTCGCGCTCG CCGTCGCGAG CAGGTTCCCG ACGCGCGGCC AGTACAGGAC 12420  
 GGGGGCGGCG AGCTGACCGC ACCAGAAGC GCGCGCACCC GCCTCGCCCG GCGTGAGCGC 12480  
 GCGACCGTC CGGCCGACC ACAGGGGCAC GCGGGTGGC CGCCACCGCG CGACCGTTC 12540  
 CGCGCTCGCC CTGGTCAGAC GCGCGCGCGC CCGGCGAGG GCGGGGAGT GGAAGGTAC 12600  
 GTCCGCTGG AGGGGCGCG CCGTCAGCCC GTCTGCGCC AGGCGCCGCG CCACCGCGCG 12660  
 CAGCGGCGCC TCCGGCCCGC ACAGCACCGT CTGACGCGGC CCGTTCACCG CGCCACCAC 12720  
 CACGTGCGCC CCCACGTACG GCGCGAGGTC GTCCACCGGC GCGGCCACGG CGAGCATGCC 12780  
 CCGGACGGC GCCGCGCGCA GCACCGCGTC CCGTTCGGCG AGCAGCCGGA GCGCGCCCC 12840  
 CCGCGCGCAC ACCCCGGCGA GCGCGGCGGC GCGAGTTGG CCGACGCTGT GCCCGAGCAA 12900  
 CAGGTGCGGC GCACCTGCG CCCCCGCGC GAGCGCCCGC CCGAGCCCGT ACCCAAGGC 12960  
 GAAGAGCAAC GGTCGCGCA CAGCGGCCCG TTCCAGCGCC TCCTGCGCG CCGCCCGTTC 13020  
 CAGCCACCGC GCGCGCACT CCTCCCCCA GGTGCCCCAC ACGGCGAAGC ACTCCTCCAC 13080  
 CCGGTCGTG AACTCGGCA CCACGCGTG GAGTCCGGCC GCCATCCGGG CCCGCTGCGC 13140  
 GCCCTGACCG GGCAGCAGCA GGGCCACCAC CCGGGCGCCC TCCTCCTCG TACCGCCCTC 13200  
 GCTCCGCACT CTTCGCTTC CTGTCCTCG CCACTCCTGC CCGTACCGAG CCTCGCCCTT 13260  
 GCCGCTGAC GCGCGTGGA CTCCTGCTCC AGCGCGGCC GCGGGGGCG CCACGAGCC 13320  
 TCAAGCCCGC TTCGAGGGAC GTCTTCCAAG GTCGGTGA GCGCGTGAAC GGCACCGCGC 13380  
 CGTCCAGGAC GCCAAGGGG GTCGTGAGT GAACCGACAA GTCGCGTCA CCGCATCGG 13440  
 CGTGGTGGCT CCGGGAGGGA TCGGGCGCAA GCCGTACTGG GAGCAGCTCA CCTCGGAG 13500  
 CACCGCCACG CGCGCCATCT CCTTCTTGA CGCCTCGCCC TTCCGCTCGG GGATCGCCGC 13560  
 CGAGGTCGAC TTCGACCCCG CCGGCGCGGG CCTGAGCCCG CCGAGGTCC GCGCATGGA 13620  
 CCGGGCGCGC CAGTTCGCG TCGTGAGCGC CAGGGAGAGC CTGCGCGACA GCGGACTCGA 13680  
 CGTCGCGAC CTCGACCCC ACCGGATCG GGTGAGCATC GGCAGCGCG TGGGCGGAC 13740  
 CACCTCGTG GAACGCGAGT ACCTCGCCT CAGCGACAGC GGGCGACAGT GGGAACCTGA 13800  
 CCTCTCTAC CTCTCGCGC ACCTCTACGA CGCCTTACC CCCAGCTCG TCGCCCGCA 13860  
 GGTGGCCGGG GTGATCGCG CCGAGGGGCC CGCGGCAGTC GTCTCCACCG GCTGCACCTC 13920  
 CGGCATCGAC TCGCTCGGC ACGCCCGCGA CCTCATCGCC GAGGGCAGCG CCGACGTCGT 13980  
 GCTCGCGGC GGCACCGACA CCCCCATCT GCCGATCGCC GTGCGCTGCT TCGACGCCAT 14040  
 CAAGGCCACC TCGCCAGCA ACGAGACCC GCGCACGCC TCGCGCCCT TCGACAGGGA 14100  
 ACGCAACGGC TTCGTCTCG CCGAGGGCGC CGCGCTCTC GTGCTGGAGG AGCTGGGGCA 14160

CGCGGGGCC CGGACGCC ACCTCTACGC TCTCGTGTCC GGCTACGCCA CGCGCTGCAA 14220  
 CGCGTACCAC ATGACCGGGC TCACCCCGCA CGGCAGGGAG ATGGCCGAGG CGATCCGGCA 14280  
 CGCGCTCGCC GAGAGCGGA CCGACCCCGC CGCCGTGGAC TACGTCAACG CCCACGGCTC 14340  
 CGGCACCAAG CAGAACGACC GGCACGAGAC CGCGCCTTC AAGGCCACGC TCGGCGAGCG 14400  
 GGCCAGGAGC GTGCCGCTCA GCTCCATCAA GTCGATGATC GGCCACTCGC TCGGTGCCAT 14460  
 CGGTTCGCTG GAGATCGCGG CCTGGGCGCT CGCCATGGAG TACGGCGTGG TGCCGCCAC 14520  
 CGCCAACCTC GACACGCCCG ACCCGAGTG CGACCTCGAC TACGTGCCCC ACGAGGCGCG 14580  
 CGAGACCGCG GTGGACAGGG TCCTGAGCGT CGGCAGCGGC TTCGGCGGTT TCCAGAGCGC 14640  
 CATGGTCTC ACCCGGGACA CGGGCGCCCG CCTCCGAGC GCTTACACC CGCTCACCGC 14700  
 ACGCCCGCA CGGCCCGGA CGGCCGGCCA GGTCCCGCAC CGCTCGACA CGGAAGGCAC 14760  
 CGGCCATGA CCACGCCCC CTCCCGACC GCCAGGGCG CCCCTCCCGG GGCCGCCCTG 14820  
 CCCCCGTCT TCACCGGCAT CGGCGTCGCC GCCCCAACG GGCTCGGCAC CGAGGAGTGG 14880  
 TGGGCGGCCA CCTGCGCGG GGAGCAGGT CTGCGGCCCG TCACGGAGTA CGACGCGAGC 14940  
 GGCCACCCCG GTGGAATCGT CGGCCGCGTC CGGACTTCG ACGCCGCCCG CCACCTGCCC 15000  
 GGCCGCTCC TGCCGAGAC CGACCGGTC ACCCGGCTCG CGCTCGTCGC CGCCGACGAG 15060  
 GCACTGAAGG ACGCCGCCGT CGACCGGCC CGGCTGCCCG AGTACGGGG GAGCGCCGTC 15120  
 ACCTCCAACG CCACCGCGG CTTCGAGTTC ACCACCGCG AGATCCGCA ACTCTGGACC 15180  
 GAGGGCCCG CCCCGCTCAG CGTCTACGAG TCCTTCGCTT GTTCTACGC CGTCAACACC 15240  
 GGCCAGATCT CCATCCGCA CGGCATGCGC GGCCCGCGG CCGTCGTCTT GGCCGACCAG 15300  
 GCGGGCGGCC TCGACGCCCT CGGCCAGGCC CGCCGCGTAC TGCGCAAGGG CGGGGTGCTG 15360  
 GCGGTGAGCG GCGCGTGGA GTCCGCGCTC GACCCCTGGG GCCTGGCCG CCACGCCTCG 15420  
 TCGGGACCC TCAGCGCTC CGCGGACCCG GCCACCGCT ACCTCCCTT CGACCGCGT 15480  
 GCCCTCGCA CCGTCGTGCG CGAGGGCGG GCCCTCCTCA CCTGGAGAC ACCCGGCAC 15540  
 GCGGAGGAG GCGACGCACC CCGATCTAC GCGAACTCG CCGGTACGC CGCCAGTTC 15600  
 GACCGCCCG CGGGTCCGG ACGCCCCC GGCCTCGAAC GCGCCGCGG CCTCGCCCTC 15660  
 GCGGACGCG GCCTGGCACC CGGGGACGTC GACGTCGTCT TCGCGGACGC GCGGGGCTC 15720  
 CCGCCCGCG ACGCCGCCGA GGCCGCGCC CTGCGCGCGC TCTTCGGCCC CGCGGCGTT 15780  
 CCGGTGAGCG TGCCGAAGAC CCAGACCGGG CGGCTCGCTT CGGGCGGCC GGCCTCGAC 15840  
 GTGCGCGCG CCTCCTCGC CTGCGCGAC GGCCTCGTC CCGCGCGT CCACCTCGAC 15900  
 GAGGTGATC CCGGTACGG CCTCGACCTC GTCCGCGACA CCGCGCGCG CCTCCGCTG 15960  
 CGCACGGCG TCCTCCTCG GCGCGGCAC GCGGCTTCA ACGCGCGCT CGTCTCCGC 16020  
 GGGCGCGGC GGCCCGTAC CGCTGAGGG CCTGCCCGG GTGCGCGCC GCGCGCGAC 16080  
 CGCGCTCCA CCGCCCGCA CACCCACAA GGAGTTCCA TGAGCGACT GACCGTCGAC 16140  
 GACCTCAAGA AACTGCTCGC CGAGACCGC GGGGAGGAG ACAGCGTGA CCTCGCCGGA 16200  
 GAACTCGACA CGCCCTTCTT GGACCTCGG TACGACTCC TCAGCTGCT GGAGACGGC 16260  
 GCGGTGCTC AGCAGCGTA CGGCATCGC CTGACTGAC AGACGGTCG CCGCTGGGC 16320  
 ACCCGCGCG AACTGCTCGA CGAGGTCAAC ACCACCCCG CCACCGCTG AGCGGGCGG 16380  
 GCACGCGAAC GAAGGTCCG GCCCCACCC CGCGCGGGG GGTGGGGCG GACCTTCGT 16440  
 TCGCGCGGA CCTCACAGCA AGGTGCGGTC CTGGGTACC TGCCGCGACA GCGTGCAG 16500  
 TTCCCGCCC TCCCGACCA GGACGTCCG TACGACGCA CTGGGACCA TCACGGGAG 16560  
 CGCCGCTCC GGCCGGTCA CCACGACCAG CGCGTAGAAC TCCGTCTGA CGCGCCGTC 16620  
 CTCACCGGG CGTACGTCCA GCATCGAGAA CCAGTGCCG CGCTGGACG GGTCTCGG 16680  
 GAACCGCGG GCGTTGAAG CGCGACCTC CTCGCGATC GCGCCCTGC CCGGGCCG 16740  
 CGGGGTGTC GGGGAGTGG CGAAGACGCC CTCGCGCGG AAGGTGGCG CGAACTCCTC 16800  
 GGCCCGTACC GCGTCCAGT GGTGTGCTG GCGCGCTAG AACGCTGTA CCTCGGTGTA 16860  
 GAGCGCGCG GAACGCGCG CGTCGCGCG CAGGACCGC GCCCTACGT CGTCCGCGG 16920  
 GGTGTGCTC GTACGGGAG CCTCCAGAA GGTGTCTGAC CTGTCCCGG AACCTCCGC 16980  
 GCGGTGGTC GACGGCCCT TGAGCCCCC TCAAGGGAG GCGCGAGGC TGGACACAG 17040  
 ACCCGCGGT CACACCGAC GCGCCCGCG CGAGGAGGA GGCCATCATG ACCACAGCAG 17100  
 CTCGCGACAC CCGCCCGGG GAGGACGCA CCACCGGGG ACCCGCTCT GTGACCGCG 17160

CGACCCGGGG CATCGGCTC GCCGTGCGG AGGCGCTCGT CGGCGCGGG TATCCCGTGG 17220  
 TGGTCTGCGC CCGCGACGCC GAGGCGGTG CGCGCACCGT CAAGGAGCTG GCAGCGGGCG 17280  
 GCGCCCGCGT CGAGGGCGTC GTGCGGACG TCACCGACGC CGCCTCCGTG CACGAACCTG 17340  
 TCGCCACCAC CGTCGCCCGC TTGCGCCCGG TCGAGGTCTT CGTCAACAAC GCGGGCGGGT 17400  
 CCGGCGGGGG AGTGACCGCC GAACTCAGCG AGTCCCTGTG GGACGACGTC ATCGCCACCA 17460  
 ACCTCAAGAG CGTCTTCCTG GTCACCCGGG AGGTGCTCAC CACGGGCGGG ATGACCGGGC 17520  
 GCGGCCGCGG CGTCGTCAAC ATCGCTCCA CCGGCGGCAA GCAGGGCGTC GTCTTCGGCG 17580  
 CCCCTACTC GCGGTCCAAG CACGGCGTGG TCGGCTTCAC CAAAGCCCTC GGCCTCGAAC 17640  
 TGGCCCGGAG CGGCATCACC GTCAACGCGG TCTGCCCGGG CTACGTGAG ACGCGATGG 17700  
 CCGCCGGAGT GCGCCGCCAC TACGCGGACC TGTGGGACGT CACCGAGGAG GACGTGCTGG 17760  
 CCGCTTCGA GCGAAGATC CCGCTCGGCC GGTACACCGG CCGCGACGAG GTCGCCGCC 17820  
 TCGTGGACTA CCTGGTCACC GACGCGCGG CCGCGGTGAC CGCCAGGCC CTCAACGTGT 17880  
 GCGGCGGACT GGGGAACCTG TGATGACCAC CACCACGACC GGGCACGAG GCCCGGGTCT 17940  
 CCGCGAACAC TCCGCGCGCC TCGCCGCCCC GCCCGCCTCC GCCTACGAAC TCGTGGCGGA 18000  
 CGTGACGCGC TGGCCCTCC TCTTACCCC GTGCTGACG GCGAGGTGC TGGAGAGCGG 18060  
 TCCCGGCACC GAACGCGTGC GGCTGTGGG CCTCACCGGC GAGCAGGTGC GCGGCTGGAC 18120  
 CTCGCGGGGC ACCCTCGACA GCGAGGGCCT GCGCGTCGGC TTCCGCCAGG AGGACAGCGC 18180  
 CCGCGCGTCT GCCGCGATGG GCGGGGAGTG GCGCTTACC GAGGAGGGCG AGGACACGCG 18240  
 CGCGTACTC GCGCAGGACT GGACGCTCAC CGAGCCGGGC GCGCACCGC ACCGCTGGGT 18300  
 CACCGAGACA CTCGACCGCA ACAGCACCGC CGAGATCGGG GCGGTGACCG CCTGGGCCGC 18360  
 GCGCACCCAC GCGGCGGGCG GCGCGGACGC GCTGCTCTC TCCTTACCG ACAGCTCGA 18420  
 CATCGCGCC CCGCGCGCCG ACGTGTACGC CTTCCTGAC GCGCGGACC AGTGCGCCGC 18480  
 GCGGCTCCCG CACGTACGCC GGGTGCCTT CTCGACCACC CCGCGACGC CGCTCACGGC 18540  
 GGGCGCGGAG GTGACGACC TGGAGATGGA GACGCGCGG GACGACGGA CCGGCCACCT 18600  
 CACCGCTCG ATCGGCTCG GCTTCGCGG CCGTCTCTC GTCTACAAG AGACCACTT 18660  
 GCGCGCCCCG CTCCTCGGC ACGCGGCTC CTGGGCGTG GAGCGCTGC CCGGCGGCGG 18720  
 CACCGGGTC ACCGCGCGC ACGGGTGGC ACTCGACCG GACGCGTGA CCGAACGCTT 18780  
 CCGCGCGGG ACCACGCTG CCGCGCGCG GGACACGGTA CCGGCGCTG TCGGCGGCAA 18840  
 CAGCAGGGC ACCCTGGAG GCGCCCGCG CCACACGAG TCGGCGGGG AGCGATGACC 18900  
 ACGACCGCC CCGGACTCC CTTGGCCCC GGGCCCGCAT CCGCGGAGC CGTCGCGCTC 18960  
 GCGGACGGC TGCACTCTG GCTCCAGCC GACGCGGCT GGTGCTCAG CAACGCGGGC 19020  
 ATCTGCTCG CGCCGACCG CGTCGCGCT GTCGACACG CCGCCACCG GCGGCGCTCC 19080  
 CCGCGCTCG GCGCGCGGT CCGCGGACT AGCCCGCAT CCGTGGACT GCTCGTCAAC 19140  
 ACCCACTTC ACGGGGACCA CAGCTTCGG AACGGCATC TCGGCAAGGA CGCCGTATC 19200  
 GTGCGCCAC AGCGACCCG CACCGAGAT GCGAGGCGG GACTCGGGT CACCGACTG 19260  
 TGGCGGGCG TGGACTGGG GCACGTGGG CCGGTCTGC CCCAGCTAC GTACCGCAGG 19320  
 CGCCTACGC TCCACCACG GGACCTCGC GTCGAGTGA TCCACCCCG CCGGCGCAC 19380  
 ACCACCAAG ACACCTGGT GTGGTCCCC GAGCAGCGG TCCTGTTGC GGGGACGTA 19440  
 CTCCTGCCG GGGCGACGC CTTCGTGCTG ATGGGCTCG TGACCGGCT GCTGCACACC 19500  
 CTGCGTCTG TCGCGCGCT CCGCCCCGC GTCGTCTCG GCGGGCACG CCCCTCGG 19560  
 GGGCCCGAG TGATCGAGG GACGAGCGC TACCTGCTG GGCTGCGCG CATCGCCACC 19620  
 GAGGGCCAC CCGCGGGCT CACCCGCTG GAGGCGGAC GCGCCACG CCCCGGGCC 19680  
 TTGCCCCACT GGAGCGAGC GGAACGACT GCGCCAATC TGACCGGGG GTACCGCAA 19740  
 CTCGGCCCCG CACCACTGG CACGCGCTC GACGTCTCG CCTGCTTGG CGACCTATC 19800  
 GCCTACAAG ACGGCGAGT GCGGTGAG CACGCTGAA CAGCCCGGC ACCGGCGCT 19860  
 CACGACCGC GCGCGGCCCT GCGCGGCTC GCGCGGCCG TGCCCGCGG AACGACCGG 19920  
 GCGGGCGCC CCCGAGGGA ACGGCGCGG CCGGTCTCC GTGCGGACG GCGGAGCAG 19980  
 CCGCGCGCT CACCCGCCA CACGCTGGG CCGTCCGGG CGGAAGCTG CCGCTGGTC 20040  
 GCGACCCAA GTGGCGAAG ACGCGCGGG CCGACGGTC ACGTGGTGA CACCGGGCT 20100  
 CACGGCGGT GGGTGTCTG TGAAGAACC CTCGACGCC GCGAGCGCC GCGGTACGG 20160

CTCGGGGATG TGCCGGGCGG TCAGCTGCAC CCACTCCTCC TCCGGCACCG GCTCCACGGG 20220  
 CAGCGGGCGG CCCAGTTCGG CGGCGAGCAC GGCGGTCCGC CCGGGCAGCG TCAGCACCTC 20280  
 GGGACCCGAG AGCGGCACCG CTTTCCGGC CCACCGCTCC ACGGGCGCGG TGAGAACGGC 20340  
 GGCGGCCACG TCGGCGATGT CCGCGGGTC GATCAGCGGC ACCGCGAGCC CGGCGTGC GG 20400  
 GACCGCACC ACACCGTCCC CGTGGAGCGC GGGCGCCAC AGCAGCGCGA GCGAGGCGAA 20460  
 GGCGCCCGGA CGCAGGAACA CCCAGGGCAC CCCGCTCCCG CGCACGTGCC GCTCGCCGCG 20520  
 GACGTGCTTC TGGAACATCG GGTGTGCTA CTCGGGTGA GCCGCCGCGG TCGCGGAGAG 20580  
 CACCACCACC TCGCGCAGCC CCGCGCCCTT CGCCGCGGCC CGCGCGAAGG GCGCGCCGGC 20640  
 CTCGTCTTGG AGCAGCAGGA ACGCCCGTTC CACGCCGTGC AGCGCCGCGG CGAGCGAAGC 20700  
 CGTGTCCGCC AGGTCCCCGA CGACGTCTC GGGCCGACC GCGGCGAGCG CCCGCCCGCG 20760  
 CGCGGCGAG CGCACCAGGG CGCGCACCGG CACCCCGAG TCGAGCAGCC GGGACACGAG 20820  
 GCGCCCCCGG ACCGCGCCGC TCGCCCGGC CACCAGCACT GTTCCCGCA GTGGTCCCGC 20880  
 CATCGTTCCT GTACCCCTTC CCGTACCGG TTCGCTTCT GTACCGCCA GGCGCCGTAG 20940  
 GCGGCGCCCG CGGCCTCGGG GGAGAGCGCG ACGTGCAGG CGCCGGTCTG CACGTCCCGC 21000  
 CAGAACCGCT GCACGGGGTC GTGCGGGCTC TGCCCGAGG TCCCGCTCGC ACGGACGAGC 21060  
 CGGTGCACCG CCCGACGGC CAACTCGGG GCGAGGGCGG GTTCCCGCTT GCCGCGCACC 21120  
 GCGGCGGGG GCTCGGAAG GAAACGGTG CCGTCGGCG CGCCCGCCG GCGCAGCAG 21180  
 AGCAGCTCGG CCGGTTCGGC CTCCCGGCC GCCCGGTCA GCTCCCGAG TGCTCCGGGC 21240  
 GAAGGCGCCC GCCCTGCGC GGGCGGCTCG GTCCACGCC GCAGGGCGCC CCGCGCCGCG 21300  
 CCGACGAGCG GCGTCACGAC GGGCAGTGCC CGACCCCGG CCATCCGCAC CGTGGGCACC 21360  
 GCGCCCGCGA CGCGGGGGCG CTGCCCGCG AGGACACGGG TGCGCGGGA GAAACGGTGC 21420  
 GCGGGGACGA GGACGTCTG GAGCAGAGG GTGTGCTGC CCGTCGCGG CATCCCGTC 21480  
 GTGTTCAGG TCCGTTCGAC GCGGGTGCC GCGCGCGCC CCGCAAGAA CCGACCTCG 21540  
 GGGCCCTCT CGCCCTCGC GCCCGCACG GTCGCGCAG CGAGCACCA GCGGCGTGC 21600  
 CGCAGCCCG TGATGTAGG CAGGCCCGG TCGAGCGCC AGCCCGCGG GACGGGCGTC 21660  
 ACGGTCCCG AGGGGACGAG GGGCCCGCG AGCAGCGGT CCGGGCCCTC GGCCACACC 21720  
 TCGGCCTGCC CCTCTCGGG CAGGTACGG GCGTACCGC CGGCGTAGG GGAGAGCGAG 21780  
 GCCGCCACG CGGCCGAGG GCAGCCCGT CCGAGGCGA GGACGGCTC GCGAGTTCC 21840  
 CGAAGCCGC CTTGCGGCC GCGTGGGCC GCGGCACGA AGTGGCGGG GAAGCCCGC 21900  
 GCCACGATG CCGCGACGA CTTCCCGGC GAGACCCCG TCCTCTCAC CGCGCGCGC 21960  
 GTGCCCGCC GCGAGCGCG CAGCGCCCC GGCCTTCC GCGAGCGGT CCGGGGGCGT 22020  
 CCGGCCCCCT CGTCCGGCG CCGGTCCGC TCGGCGTGT CCGGTGCGG GAGGGTGGTC 22080  
 GATTGCTCG TCATCGGTAC GTCACTCCA CGCGAACAT CTCGCGCT GCTGGAGGG 22140  
 CGCTCGACT CCGCTGGAC AGGCGGGCG GCGAGCCCTC AGGCGAGGA GTCCCGGCC 22200  
 GGTGCGCCG GCGTGGGC CCGGCGTAG ACCTGCGG CGGTGTGGC CTCGCGCGC 22260  
 AGCCGCGCG AGGCGGGCG GATGCGGG GCGCTCGCG CCGCCCGTG GCGGCGCGC 22320  
 TCGTCCAGT CGGCGAGCAT GGCCACCGC TCGGGTGGT CGGCGAGCG GTGCAGGCGC 22380  
 AGGGAGAGGA AGCCGGGCTG CGCGCGCAT TGGGCGCAGA TCTCTTCGAG GACCGGAGG 22440  
 AAACGGCTCT CGTCCCGTG CAGGAGCAT TGGTTGATCA TGGTACCAT GTGTCCACC 22500  
 GACACCGTCC CGCTTGAGT CCGCTGAAG AGGCACCGC GACACGCGG CGGGTGGGG 22560  
 GAGTGAGGG TACGGAGCG TCCGTTGAC GGTACGGGC GTCCGGTAC GCTTTCGGG 22620  
 GCGGCCGGT ACGCGCGGG GCCCGCGTG GGGTTTCTC ACCGCGGCC CCGGCGGTG 22680  
 CGTGCTCAG CGAGCGGTG GGGGAAGTC GGTACTCGA CCGCGAGAG GTGCTGGAC 22740  
 ACCCGTACGA CCTGGCAGG GTAGCCGAA TCGTTGCTG ACCAGAGTA GAGGATCGG 22800  
 TTGTGCGCGT CGACCTTGGT CGCCCGCGG TCCACGATG AGGCGTGGC CGAGCCGAG 22860  
 AAGTGGCTG AGACGGCGT GCGGAGGTG GTGAAGTGA TCTGGCGCG CAGCGCGAG 22920  
 GTGAGGGAGA CCTCGCGAG GTGGGTGAG ACCTCTCGC GGCTCGTCC GCGGCTCAG 22980  
 CGGAGGTTCA GGATCGGAT CGACACATG GGGACGGGA CCGGATCGA GCTGCCGTG 23040  
 ATCTTCGCCC CGAGTCCCG CAGACCTTC GCGACGGCG AGGCGGCACC CGTCTCGGT 23100  
 AGGACCATG TGAGCGGCG GACCGCCCT CGCGGTGCG ACTTGTGGT ATTGTCCAG 23160

```

AGGTTCTGGT CGTTGGTGAA GGAGTGCACC GTCTCGACGT GGCCGCTCTC GACGCCGAAC 23220
TCGTGCGCCA TCGCCTTGAG CGGCGGCACG ATGGCGTTCG TCGTGCAGGA GGCGCAGGAG 23280
ATGATCCGCT CGTCCGGCTT GATCGTCTCG TGGTTCACGC CGTGCACGAT GTTCGGCACG 23340
TCGCCCTTGC CGGGCGCGGT GAGGACGACC TTCGCGACAC CCGGGCGCAG ATGCTGCGAG 23400
AGGCCCGCGC GGTGCGGCCA GCGGCCCGTG TTGTCGACGA GGATCGCGTC GTTGATGCCG 23460
TAGGCGGTGT AGTCCACGGT CGCCGGGTCG TCGGAGTAGA TGAACCGGAT CGCGTTGCCG 23520
TTGCGGATGA GGGTGTCTTT CTCCTCGTCC ACGATGATCG TGCCCTGGAA CTGGCCGTGC 23580
ACGGAGTCGC GGCGCAGCAG TGAGGCGCGC TTGACGAGGT CCTGACCGGC GGTCTTGCGG 23640
ACCAACGACG CGCGCAGACG CAGGCCGTTG CCGGAGCGG CCTTCTCGAT GAGCAGCGG 23700
GCGAGGAGCC GCCCGATCCG CCCGAAGCCG TAGAGGACGA CGTCGCGCGG GGCGGCGCGC 23760
TCGATCTTCC GCTGCGCGGT CGCGCCCTCG ACGGGCCGCG CGGTGAATC GCGCAGCGCC 23820
AGGCCCGGT CGTCGCTCTT GTAGGTCTCG GCGAGCATGC CGATGTCGAT CTGGGAGGGG 23880
CCGAGGTCGA GCGTCGTGAG CGCCTGGAGG AACGGCATGG TCTCGGTGAC CGACAGCTCC 23940
TCGCCCCGA TCTGCGGGC GAATCGGTGG GTCTTGAGAA TCCCGACCAC CGACTTGTTT 24000
ACCAGCGAGC GGCTGTGCAG CAGCACGGTG ACGTCCCGTT GCCGCTGGAG CCTGCCGATG 24060
ACGGGGATCA TCGATTCCGC GATCTCCTCG CGGTCTTCC AGTTGGTGAA CGAGTCTCG 24120
TTGACAGTCA CAGATCCATC TTTCGAGCTA GGCGGCGCTC ATATGTTAAC CCGATGGGTG 24180
TGATCATCTG CCGGGCGCGG GTGTGATGTA CGACCAGCTC GGGGCGGGTG CGGGACGTG 24240
GGGATGGGGG CGCGGAGCAC CGGGTGCGGG GTGTGGGTGG CCGGTGGGGG GACGTGCGGC 24300
GGGGGTGCAG TACGTGCGGT GCGGTGCGG AACGTGCGGC GAGGTGCGGA ACGTGCGGTG 24360
CGGGTGCAGA ACGTGCGGC

```

## 【図面の簡単な説明】

【図1】フレノリシン遺伝子クラスターの制限地図である。

【図2】種々の遺伝子サブクラスターへの機能を割り当てた仮定的なフレノリシン生合成経路である。

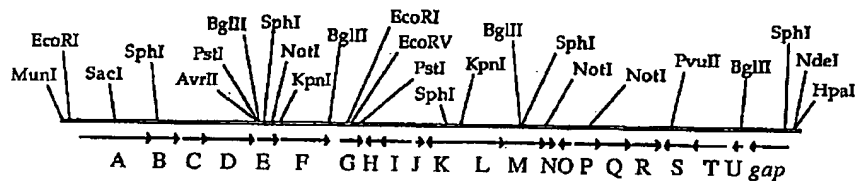
【図3】形質転換ベクターpSSVtsr およびpSSVaph の構

築を示す。

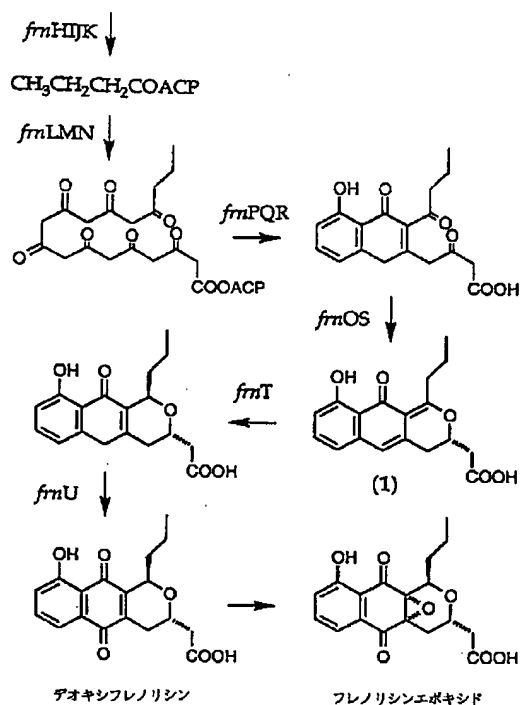
【図4】合成ermEおよびtipAプロモーターの配列を示す。

【図5】*S. roseofulvus* における遺伝子発現用ベクターの構築を示す。ここでは、転写アクチベーター遺伝子の構築物を示す。

【図1】



【図2】



【図4】

合成 *ermE*\* プロモーター:

BamHI KpnI BglII

GGATCCGGTACCAAGATCTGCTAGTTCGGGTTCATCGGGGATCGGACGGGTCGACGGGCTC

GATCTTTCACGGGTGGGAGAGCTGGGGGAGGATCTGACCGGACGGGCTGACAGCTGTC

BamHI

ACCGCGATGCTGTTGTGGGACATTCGTCGGGGTCTGCTAGGATCCagcggg

合成 *dpA* プロモーター:

BamHI NdeI KpnI BglII SmaI

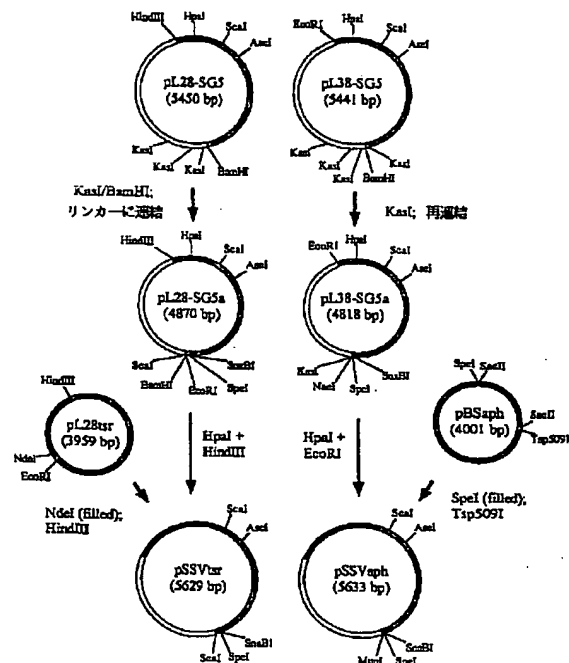
GGATCCCATATGCTGTAACAGATCTCCGGGCTGAGGGAGCGGACGGGACGGGCGGCTCA

CGGGTGGACGGGACGTCGGGGTTGCACTTCAGTCAAGTCAGGAGGACGGGCTGCA

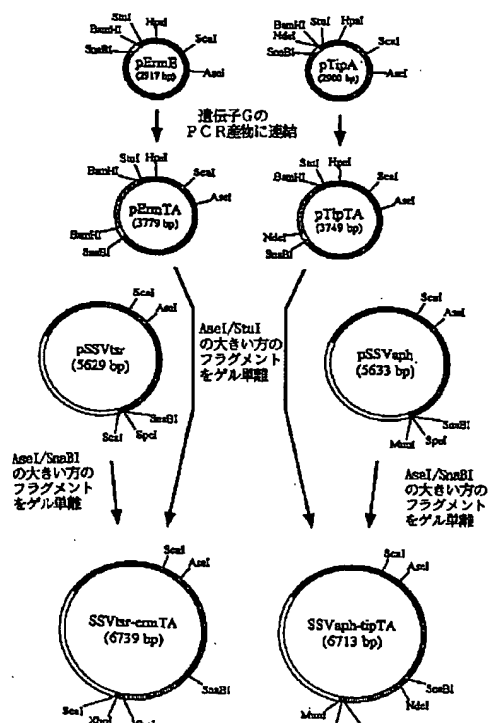
NdeI

CCGGCTCAGACAGCGGACCGCATATC

【図3】



【図5】



フロントページの続き

(51)Int.Cl. <sup>6</sup>	識別記号	F I	
C 1 2 N 1/21		C 1 2 N 1/21	
C 1 2 P 17/18		C 1 2 P 17/18	D
/(C 1 2 N 1/21			
C 1 2 R 1:465)			
(C 1 2 P 17/18			
C 1 2 R 1:465)			

(72)発明者 チャールズ・レスリー・ソリディ  
アメリカ合衆国、ワシントン 98019、デ  
ュバル、318ティエイチ・アベニュー・エ  
ヌ・イー 12627